

RESEARCH

Open Access

# Metabolic stasis in an ancient symbiosis: genome-scale metabolic networks from two *Blattabacterium cuenoti* strains, primary endosymbionts of cockroaches

Carmen Maria González-Domenech<sup>1,2</sup>, Eugeni Belda<sup>1</sup>, Rafael Patiño-Navarrete<sup>1</sup>, Andrés Moya<sup>1,3,5</sup>, Juli Peretó<sup>1,4\*</sup>, Amparo Latorre<sup>1,3,5\*</sup>

## Abstract

**Background:** Cockroaches are terrestrial insects that strikingly eliminate waste nitrogen as ammonia instead of uric acid. *Blattabacterium cuenoti* (Mercier 1906) strains Bge and Pam are the obligate primary endosymbionts of the cockroaches *Blattella germanica* and *Periplaneta americana*, respectively. The genomes of both bacterial endosymbionts have recently been sequenced, making possible a genome-scale constraint-based reconstruction of their metabolic networks. The mathematical expression of a metabolic network and the subsequent quantitative studies of phenotypic features by Flux Balance Analysis (FBA) represent an efficient functional approach to these uncultivable bacteria.

**Results:** We report the metabolic models of *Blattabacterium* strains Bge (*iCG238*) and Pam (*iCG230*), comprising 296 and 289 biochemical reactions, associated with 238 and 230 genes, and 364 and 358 metabolites, respectively. Both models reflect both the striking similarities and the singularities of these microorganisms. FBA was used to analyze the properties, potential and limits of the models, assuming some environmental constraints such as aerobic conditions and the net production of ammonia from these bacterial systems, as has been experimentally observed. In addition, *in silico* simulations with the *iCG238* model have enabled a set of carbon and nitrogen sources to be defined, which would also support a viable phenotype in terms of biomass production in the strain Pam, which lacks the first three steps of the tricarboxylic acid cycle. FBA reveals a metabolic condition that renders these enzymatic steps dispensable, thus offering a possible evolutionary explanation for their elimination. We also confirm, by computational simulations, the fragility of the metabolic networks and their host dependence.

**Conclusions:** The minimized *Blattabacterium* metabolic networks are surprisingly similar in strains Bge and Pam, after 140 million years of evolution of these endosymbionts in separate cockroach lineages. FBA performed on the reconstructed networks from the two bacteria helps to refine the functional analysis of the genomes enabling us to postulate how slightly different host metabolic contexts drove their parallel evolution.

## Background

Recently, the genomes of two different strains of *Blattabacterium cuenoti* (Mercier 1906), Bge and Pam, obligate primary endosymbionts of the cockroaches *Blattella germanica* and *Periplaneta americana*, respectively, have been sequenced [1,2]. *Blattabacterium*

constitutes a clade within the class Flavobacteria, the phylum Bacteroidetes, which contains several instances of symbionts of insects, e.g., “*Candidatus Sulcia muelleri*”, obligate endosymbiont of cicadas, spittlebugs and leafhoppers [3], “*Candidatus Cardinium*”, symbiont of the white fly *Bemisia tabaci* [4], and “*Candidatus Vestibaculum illigatum*”, which establishes a symbiosis with the gut flagellate *Staurojoenina sp.* associated to the termite *Neotermes cubanus* [5]. All these endosymbiont bacteria are relatively distant from free-living members

\* Correspondence: juli.pereto@uv.es; amparo.latorre@uv.es

<sup>1</sup>Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, P.O. Box 22085, E-46071, València, Spain

Full list of author information is available at the end of the article

within the phylum Bacteroidetes [6]. Thus, if we assume that the age of a symbiotic association of a primary endosymbiont corresponds to the oldest fossil record of its host, we estimate the time of divergence between *B. cuenoti* and its free-living cousins to be 250 Myr [7], thus being possibly one of the most ancient mutualistic insect symbioses described so far.

Cockroaches, natural hosts of *Blattabacterium* sp., excrete waste nitrogen as ammonia [8-11] unlike most terrestrial insects, which eliminate it as uric acid [11]. Ammonotelically, the ancestral character present in aquatic animals, has traditionally been considered maladaptive in terrestrial animals [12]. The enigmatic return of cockroaches to ammonotelicity seems to be related to the role of bacterial endosymbiosis in their nitrogen economy. López-Sánchez et al. [1] showed the presence of urease activity in endosymbiont-enriched extracts of the cockroaches *B. germanica* and *P. americana*. Stoichiometric analysis of the core of the reconstructed metabolic networks would suggest that these endosymbiotic bacteria participate in the nitrogen metabolism of the host. Physiological studies ([1,8] and references therein) suggest that uric acid may represent a form of nitrogen storage in cockroaches and that *B. cuenoti* may produce ammonia from uric-derived metabolites provided by the host. In fact, the cockroach fat body contains specialized cells storing uric acid (urocytes) that are in close proximity to the cells containing endosymbionts (bacteriocytes) [13].

A common feature of genomes from bacterial endosymbionts is their strict conservation of gene order and remarkable differential gene losses in the different lineages [14-16]. In the case of the Bge and Pam strains, comparative genomics reveals both a high degree of conservation in their chromosomal architecture and in the gene repertoires (accounting for a total of 627 and 619 genes in Bge and Pam, respectively) despite the low sequence similarity observed (~85% nucleotide sequence identity) [6]. Thus, the metabolic networks of these endosymbionts should be similar, differing only slightly. These differences might be analyzed from a qualitative point of view by comparison between the inferred metabolic maps, but this approach does not allow quantitative evaluation of how these inequalities might affect the functional capabilities of each microorganism. Constraint-based models of metabolic networks represent an efficient framework for a quantitative understanding of microbial physiology [17]. In fact, computational simulations with constraint-based models are approaches that help to predict cellular phenotypes given particular environmental conditions, with a high correspondence between experimental results and predictions [18-20]. It is worth mentioning that they are especially suitable for reconstructed networks from uncultivable

microorganism, as it is the case of primary endosymbionts. Thus, Flux Balance Analysis (FBA) is one of these useful techniques for the study of obligate intracellular bacteria, since it reconstructs fluxes through a network requiring neither kinetic parameters nor other detailed information on enzymes [17]. This modeling method is based on the stoichiometric coefficients of each reaction and the assumption of the system at steady-state [21]. FBA calculates metabolites fluxes through the metabolic reactions that optimize an objective function –usually biomass production–, i.e., how much each reaction contributes to the phenotype desired.

In this study, we have reconstructed the metabolic networks of Bge and Pam strains of *B. cuenoti*, focusing on their metabolic abilities and relating them to the symbiotic interaction with their host. Following the protocol proposed by Thiele and Palsson [22], we have quantitatively predicted their biochemical potential by FBA, assuming biomass formation as objective function. In addition, in some simulations we have imposed the constraint of ammonia release from both endosymbionts, in coherence with the physiological observations [8] and as expected by the measured urease activity and the stoichiometric analysis performed by López-Sánchez et al. [1]. We have performed sensitivity and robustness analyses and deduced how these endosymbionts may be related to their cockroach hosts metabolically. We offer an overview of the remarkably stable metabolic relationships in these old symbioses as well as providing an explanation for a possible environmental cause of the loss of genes coding for enzymes in a central pathway, such as the TCA cycle in one of the endosymbionts.

## Results

### Metabolic models and FBA simulations

Gene to protein to reaction (GPR) associations were included in the model *iCG238*, corresponding to the reconstructed metabolic network from *B. cuenoti* Bge strain. This model accounted for 238 genes with a known locus in the genome, linked to 296 GPR associations and with 364 associated metabolites. The model *iCG230* of the reconstructed network of the *B. cuenoti* Pam strain comprised 289 GPR associations, with the participation of 230 genes and 358 metabolites (see Table 1 and Additional Files 1 and 2). Both models included 47 exchange reactions. A difference between the two models deals with the simulated uptake of the sulfur source. Thus, due to the lack of *cysN*, *cysD* and *cysI* genes related to cysteine metabolism in the strain Pam, this model simulates the income of hydrogen sulfide (H<sub>2</sub>S) instead of sulfate, as it is the case in the strain Bge. Although *cysH* and *cysJ* genes are present in the genome of the strain Pam, they represent isolated

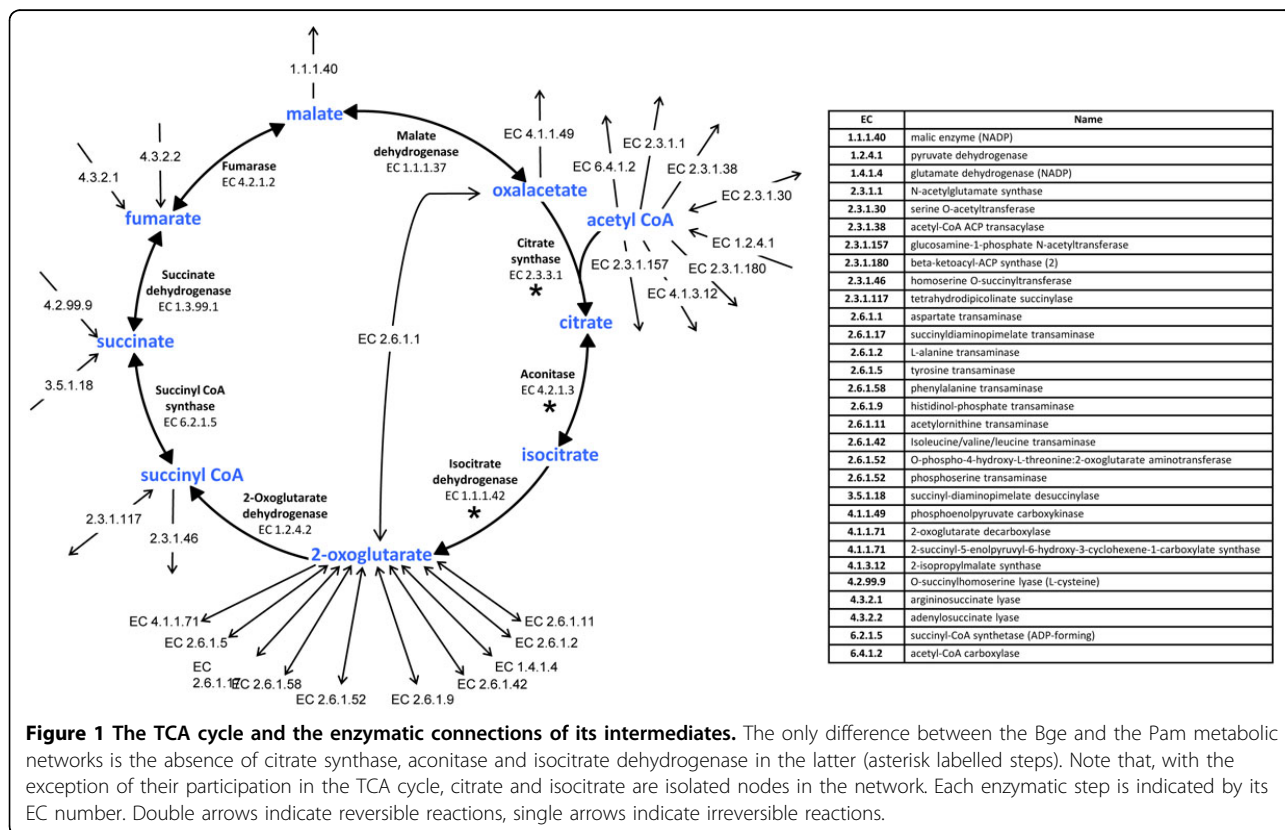
**Table 1 Characteristics of metabolic reconstructions from the strains Bge and Pam of *B. cenozi*.**

	Metabolic model	
	iCG238	iCG230
<b>Protein-encoding genes</b>	238	230
<b>Metabolites</b>	364	358
Intracellular metabolites	317	311
Extracellular metabolites	47	47
<b>Reactions</b>	418	411
Enzymatic reactions	325	318
Transport fluxes	46	46
Exchange reactions	47	47
<b>Reactions with protein-encoding gene model assignments (GPRs)</b>	296	289
Enzymatic reactions	283	276
Transport fluxes	13	13

genes within the first steps of the mentioned pathway (see Additional File 3). As a consequence, the following reactions were removed from the final metabolic network: phosphoadenylyl-sulfate reductase (thioredoxin) (EC 1.8.4.8) and sulfite reductase (NADPH) (EC 1.8.1.2), catalyzed by CysH enzyme and by the protein complex CysII (CysJ) requires the participation of CysI, also missing), respectively.

Another difference between the Bge and the Pam strain networks is the absence in the latter of the first three steps in the TCA cycle [2]. Missing metabolic steps can generate dead-end metabolites or unconnected nodes in a metabolic network. In other words, an isolated substrate (or product) is generated if it can only be consumed (or produced) by enzymes that are absent in the network [23]. However, we realized that the metabolites leading to citrate (oxaloacetate and acetyl CoA) or the metabolites derived from isocitrate (2-oxoglutarate, coenzymes excluded) are well-connected nodes in both reconstructed networks (Fig. 1), in spite of the absence of the first three steps in the TCA cycle in the strain Pam [2]. On the other hand, both metabolic models showed exactly the same 12 dead-end metabolites (see Additional Files 1 and 2). The reactions leading up to the dead ends were included to obtain a fully functional network. Furthermore, we have considered 75 reactions (33 of them being transport reactions) without any gene associated in either model (Additional Files 1 and 2, and Additional File 4 for further details).

In order to evaluate the functional phenotype of the metabolic networks from both strains, FBA with biomass production as objective function was employed, using as a reference model the reconstructed network and biomass equation of *E. coli* with some adaptations,



as described in Methods. Non-essential amino acids L-Asn, L-Gln, Gly and L-Pro, as well as the compounds (S)-dihydroorotate, nicotinic acid, pantotheine-4-phosphate, porphobilinogen and thiamin were supposed to be supplied by the host to meet the biosynthetic needs in both strains, as suggested by the genetic lack of the corresponding synthetic machineries [1,2]. The rest of essential components of the extracellular medium were CO<sub>2</sub>, Fe<sup>2+</sup>, H<sup>+</sup>, H<sub>2</sub>O, K<sup>+</sup>, Na<sup>+</sup>, O<sub>2</sub>, P<sub>i</sub> and the appropriate sulfur source (Fig. 2). All the above-mentioned chemical components of the environment (host) were necessary and sufficient to yield a viable phenotype in FBA simulations with the *iCG238* Bge strain model (Fig. 3). However, with the Pam network we obtained a mere 20% of the biomass produced by the Bge network under the same minimal conditions (Fig. 3).

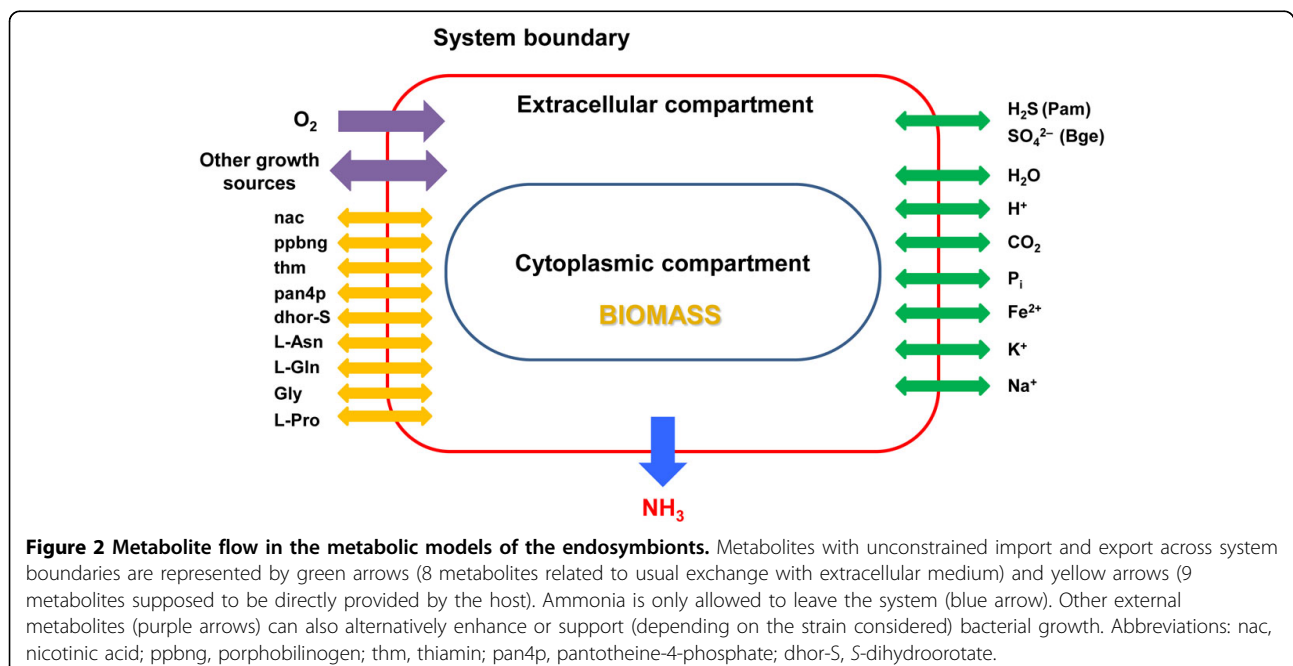
FBA was also used to predict the behavior of the strain Bge in terms of growth rate when an additional metabolite was considered in the medium. We tested several metabolites with transport systems encoded by genes present in both *B. cuenoti* genomes (L-Asp, D-fructose, fumarate, L-Glu, glycerol, L-malate, succinate and urea) and also the input of the TCA cycle intermediate 2-oxoglutarate, as a simulation of an anaplerotic reaction. All the considered additions had a positive effect on the biomass production rate by the Bge network, compared to the minimal medium (Fig. 3). In particular, some intermediates of the TCA cycle improved the performance of both networks with a remarkable ten-fold enhancement of biomass production by the Pam network in the presence of L-Glu and 2-

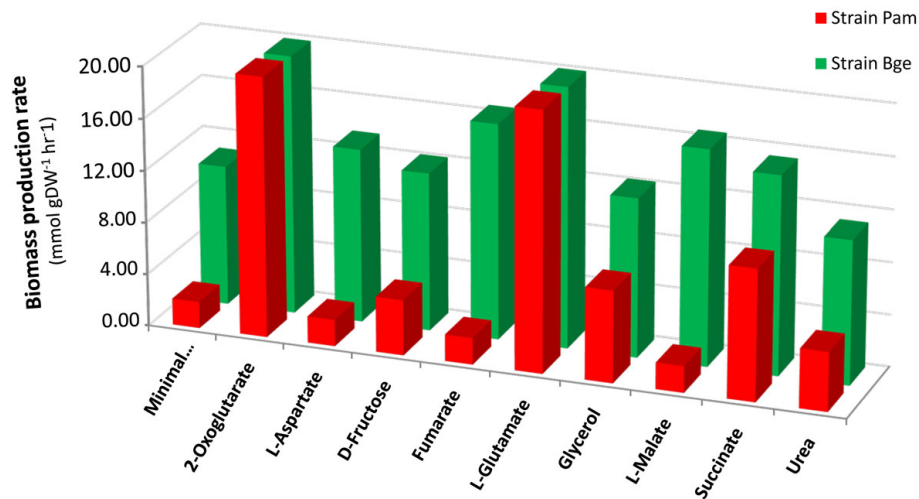
oxoglutarate. This result can be correlated with the fact that the strain Pam possesses an incomplete TCA cycle. We decided to focus our attention on how the metabolic flux should be completely redirected through the different reactions leaving or entering this pathway (see Fig. 1). Thus, the fluxes through the transaminase reactions generating 2-oxoglutarate were particularly important because they feed the enzymatic steps of the TCA cycle downstream of the isocitrate dehydrogenase reaction (Fig. 4). In summary, the positive effect of L-Glu (and 2-oxoglutarate) on the Pam network achieved a similar performance to the Bge network due to the anaplerotic effect of these metabolites (Fig. 4).

The excretion of ammonia from the system, a phenomenon compatible with the physiological and experimental observations (for review see [8] and [1] and references therein), was always observed in simulations with both models under minimal conditions. The efflux of ammonia reached maximum levels when L-Glu uptake was simulated by the system. However, the efflux of ammonia stopped and could even be reversed when 2-oxoglutarate or succinate were provided to both metabolic networks. This was due to an increased assimilation of ammonia through displacement of the glutamate dehydrogenase reaction (EC 1.4.1.4) in the assimilative direction.

#### Sensitivity and robustness analysis

The robustness of both metabolic reconstructions, *iCG238* and *iCG230*, was explored from different perspectives. Firstly, we performed a sensitivity analysis, i.e.

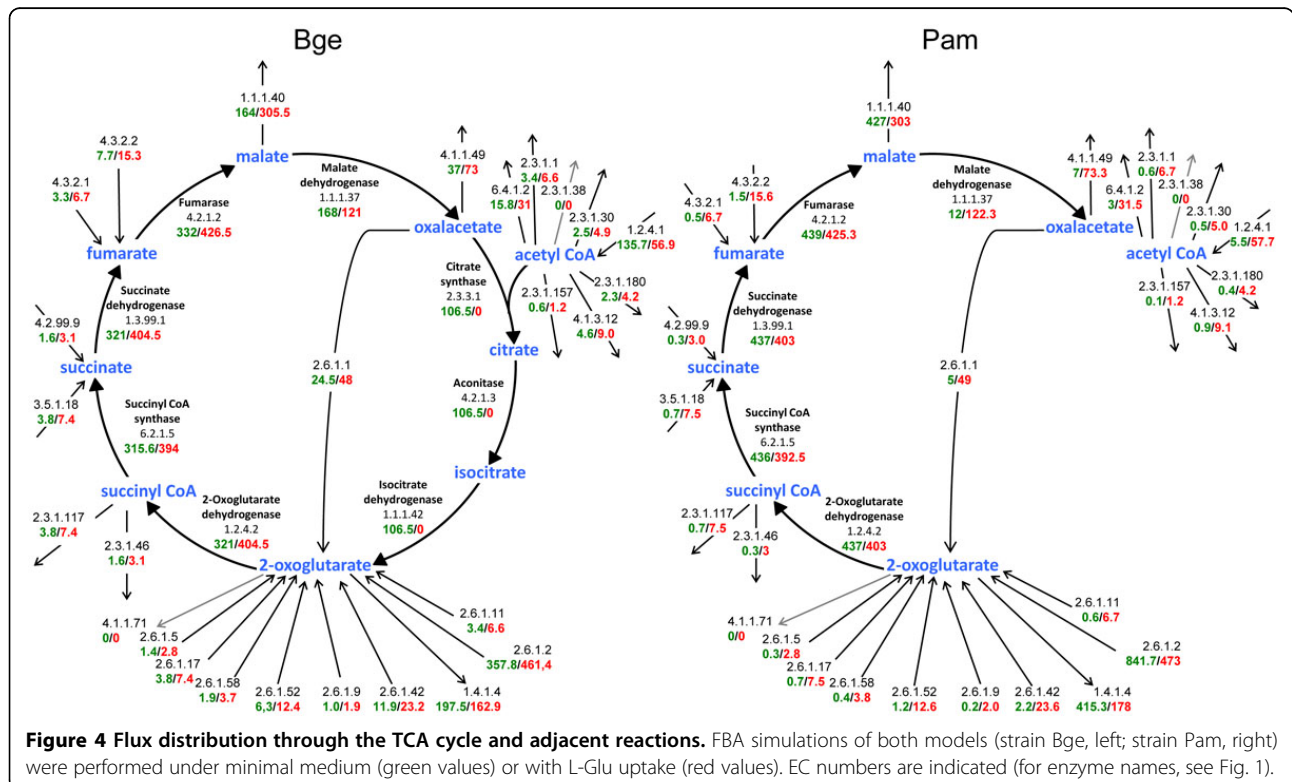




**Figure 3 Effect of different metabolites on the performance of the metabolic models.** Biomass production rates (mmol g DW<sup>-1</sup> h<sup>-1</sup>) in the two networks (strain Bge, green bars; strain Pam, red bars) were measured under minimal conditions (see Fig. 2 and main text) or considering the uptake of different metabolites.

how biomass production rate changed as the flux over a specific reaction of interest varied in magnitude. The target reactions to perform this analysis were those involving the exchange of essential and additional growth sources used in the FBA simulations described in the previous section. We also analyzed the effect of oxygen uptake since the metabolic inference from the

two cockroach endosymbiont genomes indicates the presence of a complete electron transport chain terminated with a high-affinity *cbb<sub>3</sub>*-type cytochrome oxidase [1,2]. Furthermore, the cockroach fat body, the tissue where endosymbionts are located, exhibits the characteristics of an active aerobic environment (e.g. peroxisome abundance and urate catabolism, [23,1] and



**Figure 4 Flux distribution through the TCA cycle and adjacent reactions.** FBA simulations of both models (strain Bge, left; strain Pam, right) were performed under minimal medium (green values) or with L-Glu uptake (red values). EC numbers are indicated (for enzyme names, see Fig. 1).

references therein). Both the *iCG238* and the *iCG230* models, showed a strict dependence on the import of L-Asn, Gly and L-Pro, in accordance with the metabolic inference from the genomes [1,2]. Our simulations using Bge model show that there is a range of metabolic flux values for oxygen and L-Gln exchange reactions over which it is possible to produce an optimum phenotype in terms of biomass (Fig. 5). A similar result was observed for the growth dependence on L-Gln with the Pam model (data not shown).

We also evaluated the sensitivity of the Bge metabolic network to variations in the three first reactions of the TCA cycle, absent in the metabolic network of the strain Pam ([2]; see Fig. 1). We simulated the minimal conditions and those considering the additional uptake of some intermediates of the cycle as well as the anaplerotic amino acids L-Glu and L-Asp, precursors of 2-oxoglutarate and oxalacetate, respectively. As shown in Figure 6, a viable phenotype is produced even when the flux values through the three aforementioned reactions are null. Moreover, the biomass production reaches a maximum value when the flux across such reactions is zero and 2-oxoglutarate or L-Glu is added.

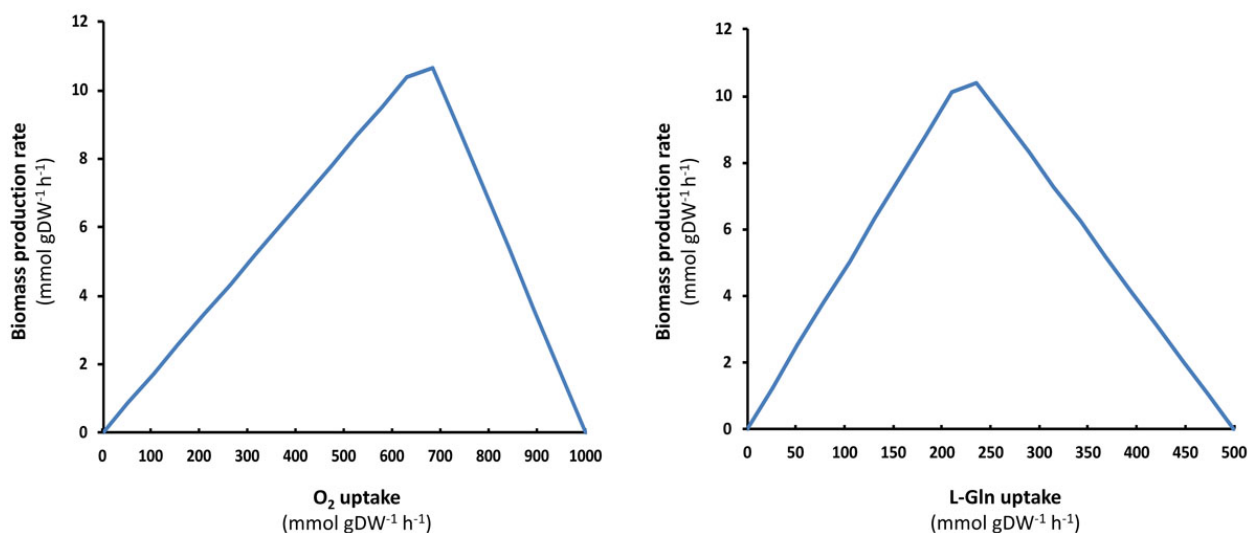
Finally, we also explored the robustness of both metabolic networks by randomly removing genes. The starting point was the aerobic and minimal conditions already described and then we also evaluated the fragility of the network when additional carbon sources were provided to the system. As cutoff value for gene essentiality a >99% decrease in the biomass production after the gene deletion was used, as described by Thomas et al. [24]. For the Bge strain network, a set of essential genes was determined ranging between 76.1 % (minimal

medium) and 72.3 % (with added glycerol) of the total genes comprised in the model. With the Pam network we found a genetic essentiality between 79.6 % (minimal medium) and 73.5 % (with added fumarate, L-malate or glycerol).

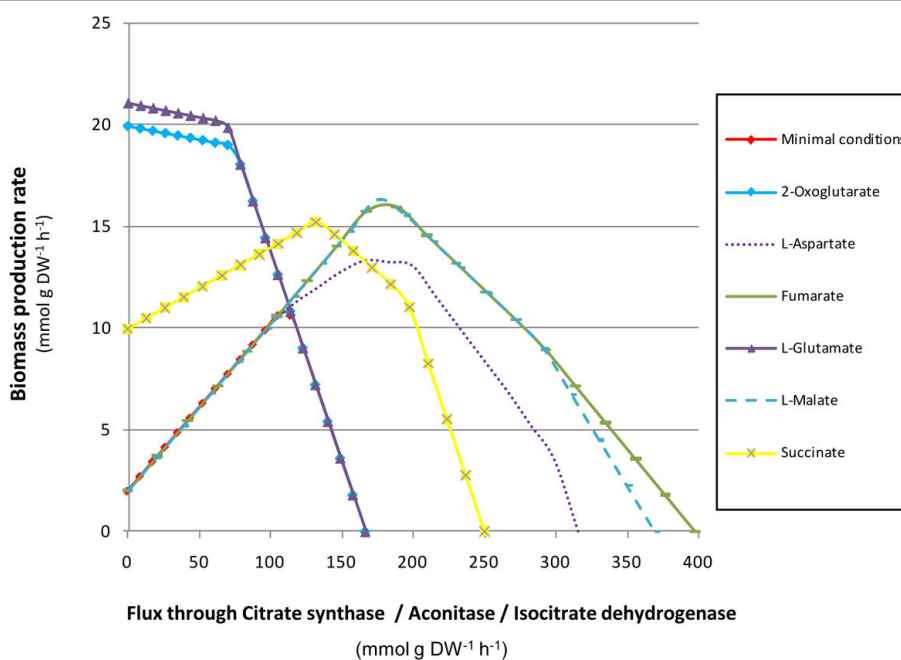
## Discussion

### Uncultivable bacteria can be studied by *in silico* simulations

In this paper we describe the genome-scale metabolic networks corresponding to two strains of *B. cuenoti*, Bge and Pam, the endosymbiotic bacteria of the cockroaches *B. germanica* and *P. americana*, respectively. Despite the approximately 140-Myr of parallel evolution, both metabolic networks showed striking conservation and we decided to compare their functionality by means of a stoichiometric approach such as FBA. This computational methodology has already been successfully used in a study of the metabolic network robustness of *B. aphidicola*, the primary endosymbiont of aphids, in comparison to *E. coli* [24] and for the simulation of reductive evolution in endosymbionts [25,26]. Thus, FBA represents a valid strategy for the functional study of those bacterial species that pose important obstacles to their empirical study, as it is the case of the uncultivable endosymbionts. In this work we used the *E. coli* model as a reference since to the best of our knowledge there are no empirical data on the biomass function of any members within the phylum Bacteroidetes. In the absence of information related to real biomass composition of the modeled organism, the use of the equations of *E. coli* is considered a reliable approach and an acceptable starting point [19,27-29].



**Figure 5** Effect of oxygen and L-Gln uptake on metabolic network performance. Biomass production rates (mmol g DW<sup>-1</sup> h<sup>-1</sup>) by the Bge strain model were measured at different uptake rates of oxygen (left) and L-Gln (right).



**Figure 6 Sensitivity analysis for the first three reactions of the TCA cycle.** Biomass production rates ( $\text{mmol g DW}^{-1} \text{h}^{-1}$ ) by the Bge strain model were measured under different metabolic environments (minimal conditions or the uptake of the indicated metabolites, see inset) and diverse reaction flux through the first enzymatic steps of the TCA cycle: citrate synthase, aconitase and isocitrate dehydrogenase.

The simulations allowed us to identify the minimal environmental components for a functional metabolic network (Fig. 2). For instance, both *Blattabacterium* networks show a strict dependence on L-Gln supply from the host due to the absence of glutamine synthase in both endosymbionts. This dependence of the functionality on the availability of some chemical species may also suggest a possible regulatory role of the external medium in the metabolic behavior of the bacterium. In other biological systems, like the nitrogen-fixing nodules of Leguminosae, oxygen availability modulation by the host has been suggested as a mechanism of punishment to cheaters in the symbiotic relationship [30]. Our *in silico* simulations (Fig. 5) suggest that access to L-Gln and/or oxygen is a good candidate for a control mechanism of cockroaches over their endosymbiotic population.

#### Why *Blattabacterium* strain Pam has lost the first three enzymatic steps of the TCA cycle

One of the few differences between the metabolic networks inferred for the Bge and Pam strains of *Blattabacterium* is that the latter has an incomplete TCA cycle (compare [1] and [2]). Sabree and coworkers [6] hypothesized that the two missing enzymes, aconitase (EC 4.2.1.3, *acnA*) and isocitrate dehydrogenase (EC 1.1.1.42, *icd*), in the Pam strain metabolic network, can be functionally substituted by the enzymes 3-

isopropylmalate isomerase (EC 4.2.1.33, *leuC*) and 3-isopropylmalate dehydrogenase (EC 1.1.1.85, *leuB*), respectively. However, the first enzymatic step of the TCA cycle (citrate synthase, EC 2.3.3.1, *gltA*) is also absent and apparently there is no other alternative solution to this absent activity. Although the functional substitution of two out of three missing metabolic steps in the TCA cycle cannot be excluded, here we have shown the dispensability of all three genes to obtain a functional phenotype in terms of biomass production under certain conditions. Thus, the proposal of functional substitutions by homologous enzymes is an unnecessary conjecture in this case. There are two reasons: (i) as shown in Figure 1, the lack of the three afore-mentioned steps does not generate true dead-end metabolites, and (ii) there is an alternative way to keep a fully functional metabolic network without the first three enzymes in the TCA cycle.

Our simulations show that the Pam network behaves like the Bge network if an anaplerotic reaction (i.e. the uptake of L-Glu or 2-oxoglutarate) is provided. Under these circumstances, the metabolic fluxes are redirected around the TCA cycle (Fig. 4) and, as shown in Figure 6, the sensitivity analysis demonstrates that the flux through the first three enzymatic steps of the TCA cycle can be null. This behavior may explain the dispensability of the corresponding *gltA*, *acnA*, and *icd* genes if the host provides the endosymbiont with any of the above-

mentioned compounds. In other words, the provision of a non-essential amino acid to the endosymbiont by the host may offer a set of biochemical conditions favoring the loss of central metabolic genes in one particular evolutionary lineage.

The loss of these three enzymatic steps in the Pam strain of *Blattabacterium* is an example of how the essentiality of genes may change when the environmental conditions change. Studies of flux connectivity (i.e. reactions that always work together) [31] and synthetic lethality analysis (i.e. searching the effect of multiple gene deletions) [32] have shown that in free-living bacteria, such as *E. coli* or *Helicobacter pylori*, the enzymes coded by the *gltA*, *acnA* and *icd* genes form a subset of essential steps. This enzymatic subset was also determined during our analysis of elementary flux modes in *Blattabacterium* Bge [1]. Thus, it is conceivable that during the transition to intracellular lifestyle, the ancestor of *Blattabacterium* strain Pam found a set of chemical conditions in the host cell making those three formerly essential genes dispensable and thus allowing their loss *en bloc*.

#### **Blattabacterium metabolic networks are fragile**

The percentage of essential genes predicted for both *Blattabacterium* strains in gene deletion simulations (ranging from 72.3% to 79.6%) is lower to that determined for *B. aphidicola*, primary endosymbiont of aphids, which showed a fraction of 84% of essential genes in a similar simulation [24]. Those values of genetic essentiality in endosymbiotic metabolic networks are far from the robustness observed in models of free-living bacteria, e.g., around 15% of essential genes coding for metabolic enzymes in *E. coli* [33]. Thus, endosymbiotic metabolic networks are less redundant than networks from free-living bacteria. In comparison to the extreme fragility of a minimalist metabolic network, theoretically deduced from comparative genomics [34] and analyzed by Gabaldón *et al.* [35], with 98% of essential genes, endosymbiont metabolic networks show an intermediate degree of robustness, and may represent different stages of the reductive evolutionary process associated to intracellular lifestyle.

#### **Blattabacterium has a key role in the nitrogen economy of cockroaches**

Our working hypothesis is that *Blattabacterium* played a key role during the transition from uricotelous to a use of urates as nitrogen storage in cockroaches. The elementary flux mode analysis and the enzymatic assays performed by López-Sánchez *et al.* [1] indicated that the central metabolism of *Blattabacterium* can use urea (and some other nitrogen compounds, as non-essential amino acids) and excrete ammonia. As shown in this work, under minimal conditions the reconstructed metabolic networks of the Bge and Pam strains produce

ammonia when biomass growth is optimized. This metabolic performance is compatible with the classical physiological observations made by Cochran and coworkers [8]. In addition, physiological studies with cockroaches indicate that uric acid is a form of nitrogen storage instead of a major waste product like in most insects [8]. According to our hypothesis, the fat body metabolism would produce urea from uric acid and the endosymbiont urease would transform urea into ammonia to be used again, partially by the endosymbiont (i.e. synthesis of Glu via the displacement of the Glu dehydrogenase reaction) and partially by the host, especially for glutamine biosynthesis by Gln synthase. It is remarkable that this enzymatic reaction is absent in *Blattabacterium*, although the metabolic networks of both Bge and Pam strains contain 9 Gln-consuming reactions (in addition to the requirement of Gln for protein synthesis represented by the corresponding tRNA for Gln and a gene coding for glutamine tRNA ligase, *glnS*).

In that context, the retention of a urease in *Blattabacterium* makes evolutionary sense as a key piece of the metabolic mosaic of the cockroach nitrogen economy, whereas the bacterial dependence on a Gln supply by the host contributes to the obligate character of this symbiotic association. The dependence on host-supplied Gln has also been recently described in *Blochmannia vafer*, the primary endosymbiont of the ant *Camponotus vafer*, which contains urease but lacks both Glu dehydrogenase and Gln synthase [36].

The released ammonia observed by physiologists would correspond to the escape of some ammonia produced by the system when all the ammonia-utilizing reactions are saturated, a side effect of the serial transformation from uric acid to urea to ammonia to glutamate/glutamine. In this metabolic framework, our *in silico* modeling was performed with the constraint of ammonia release by the endosymbiont. The mathematical expression of the metabolic networks, thus, helps us understand the systemic properties of the host-endosymbiont relationships. Practically speaking, it serves for the better design of an experimental strategy to functionally characterize the pathway from uric acid to glutamine in cockroaches.

#### **Conclusions**

One of our aims was to perform a genome-scale constraint-based modeling of the metabolisms of two different strains of *B. cuenoti*, Bge and Pam, primary endosymbionts of the cockroaches *B. germanica* and *P. americana*, respectively, which are the result of a parallel evolution during the last 140 million years. A striking feature of the two bacteria is not only the genome architecture conservation, as observed in other similar systems, but also the few gene losses undergone in the



different lineages. Thus, both metabolic networks differ from each other in only seven enzymatic reactions. The FBA approach has allowed us to evaluate the different host influences that might explain the loss or retention of certain genes, which is not easy to elucidate *a priori* by visual inspection of the respective metabolic maps. In addition, the fragility shown by the metabolic networks is compatible with a constancy of environmental conditions, and it is the expected outcome for minimal metabolisms derived from the streamlining of endosymbiotic bacterial genomes. The model predictions will allow us to address future functional analyses, and formulate new hypotheses on the metabolic interdependence in the ancient symbiosis between *B. cuenoti* and cockroaches.

## Methods

### Definition of the iCG238 and iCG230 models and FBA simulations

We reconstructed the *iCG238* and *iCG230* networks using the *E. coli* K-12 *iJR904* model as a starting point [37]. From this model, we proceeded as Thomas *et al.* [24] removing all reactions associated with pseudogenes, genes without homologs in those strains or unconnected with the biomass reaction (e.g., *gltX*, *dna*, encoding genes of tRNA ligases and DNA glycosylases). We employed the OrthoMCL algorithm [38] to search for orthologs between *E. coli* K-12 and the different strains of *Blattabacterium* sp. as well as between the two *Blattabacterium* strains in order to obtain a first draft of the metabolic models (inflation thresholds, between 1.2 and 5, choosing in each case the best, normally 1.5 and 3). In addition, the urease reaction, supported by experimental evidence [1], was also added to both models, and ubiquinone was replaced by menaquinone in each metabolic reconstruction since this is the unique or majority quinone in the *Flavobacteriaceae* family [39]. To further curate the models, we performed additional BLAST searches [40] among the corresponding strain of *Blattabacterium*, other flavobacteria and *E. coli* K-12 available in GenBank (e-values below  $e^{-11}$ ), to incorporate reactions either absent in *E. coli* or undetected due to the divergence among strains. In addition, we identified functional domains by means of the interface SMART (Simple Modular Architecture Research Tool) ([http://smart.emblheidelberg.de/help/smart\\_about.shtml](http://smart.emblheidelberg.de/help/smart_about.shtml)) [41,42]. Flux balance analysis (FBA) was performed using the COBRA toolbox [43], a freely available Matlab toolbox and the models were described using the Systems Biology Markup Language (SBML) [44] (Additional Files 5 and 6).

We used the biomass equation derived from the *iJR904 E. coli* model [37] with a few adaptations derived on updated network of such microorganism, i.e.

*iAF1260* [33]. In particular we added the cofactors thiamine diphosphate and tetrahydrofolate. Additionally, we adjusted the amounts of the four different deoxynucleotide triphosphates in the biomass equation to reflect the GC content of the *Blattabacterium* strains (Bge, 27 mol %; Pam, 28 mol%). Furthermore, since *Blattabacterium* strains are unable to completely synthesize cardiolipin, glycogen, lipopolysaccharide, and spermidine, we removed these components from the biomass equation.

### Robustness analysis

The study of network robustness was performed with the function *robustnessAnalysis* of the COBRA toolbox [43]. In addition, we evaluated the effect of a gene deletion experiment on cellular growth of the resultant mutant using the option *singleGeneDeletion* of the COBRA toolbox. We set to zero the upper and lower flux bounds for the reaction(s) corresponding to the simulated deleted gene. If a single gene is associated with multiple reactions, the deletion of that gene will result in the removal of all associated reactions. On the contrary, a reaction that can be catalyzed by multiple non-interacting gene products will not be removed in a single gene deletion. The possible results of a single deletion are unchanged maximal growth (non-lethal), reduced maximal growth or no growth (lethal). We simulated growth and subsequent fragility analysis with all the different sources which enhance/support biomass formation.

### Additional material

**Additional file 1: Description of the metabolic model of the Bge strain of *B. cuenoti* (host *Blattella germanica*), containing: a list of the GPR associations; a list of the reactions that were supposed to be placed although without any associated gene; a list of the exchange fluxes used in simulations and their constraints; a list of definitions of the metabolite abbreviations; and a list of the dead-end metabolites in the metabolic network.**

**Additional file 2: Description of the metabolic model of the Pam strain of *B. cuenoti* (host *Periplaneta americana*), containing: the same kind of information as Additional file 1.**

**Additional file 3: Differences in the cysteine biosynthesis pathway between the strains Bge and Pam.** Sulfate constitutes the sulfur donor in the strain Bge, whereas this function is performed by hydrogen sulfide in the strain Pam. In green, genes exclusively present in *B. cuenoti* (strain Bge); in blue, genes extant in both bacterial strains, Bge and Pam. For all the compounds shown, see the list of abbreviations in the corresponding Metabolites section of Additional files 1 and 2.

**Additional file 4: Further details on the reconstruction of the networks**

**Additional file 5: Metabolic network model of Bge strain in Systems Biology Markup Language (sbml) format [44], ready to perform simulations with COBRA toolbox [43].**

**Additional file 6: Metabolic network model of Pam strain in Systems Biology Markup Language (sbml) format [44], ready to perform simulations with COBRA toolbox [43].**

#### List of abbreviations used

Bge: *Blattabacterium cuenoti* strain of cockroach *Blattella germanica*; DW: Dry Weight; FBA: Flux Balance Analysis; GPR: Gene to protein to reaction association; Pam: *Blattabacterium cuenoti* strain of cockroach *Periplaneta Americana*; TCA cycle: Tricarboxylic acid cycle.

#### Authors' contributions

CMGD performed the reconstruction process, analyzed the data and evaluated the models, also writing the first draft of the manuscript; EB helped actively in the analyses with COBRA and in drafting the manuscript; RPN helped in the comparative functional analyses between both strains and in drafting the manuscript; AM conceived the study and made important contributions to draft the manuscript; JP conceived and supervised the study and wrote the final manuscript; AL conceived the study and wrote the final manuscript. All authors read and approved the final manuscript.

#### Authors' information

CMGD: postdoctoral specialist in Microbiology and Systems Biology; EB: postdoctoral specialist in Bioinformatics, Evolutionary Genomics and Systems Biology; RPN: PhD student specialist in Genetics, 'omics' Sciences and Bioinformatics; AM: Full Professor of Genetics; JP: Associate Professor of Biochemistry and Molecular Biology; AL: Full Professor of Genetics.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

Financial support was provided by grants BFU2009-12895-C02-01/BMC (Ministerio de Ciencia e Innovación, Spain) to AL and Prometeo Program (Generalitat Valenciana) to AM. Dr. González-Domenech was supported by grant from the University of Granada. R. Patiño-Navarrete was recipient of a fellowship from Ministerio de Educación y Ciencia, Spain. We also thank to Mr. Alejandro Manzano for his assistance with bioinformatic issues, Dr. Alex Neef for helpful discussions as well as two anonymous reviewers for their valuable comments. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This article has been published as part of *BMC Microbiology* Volume 11 Supplement 1, 2012: Arthropod symbioses: from fundamental studies to pest and disease management. The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2180/12?issue=S1>.

#### Author details

<sup>1</sup>Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, P.O. Box 22085, E-46071, València, Spain. <sup>2</sup>Faculty of Pharmacy, University of Granada. Campus of Cartuja, E-18071. Granada, Spain. <sup>3</sup>Departament de Genètica, Universitat de València, Spain. <sup>4</sup>Departament de Bioquímica i Biologia Molecular, Universitat de València, Spain. <sup>5</sup>Centre for Public Health Research (CSISP), E-46020. València, Spain.

Published: 18 January 2012

#### References

- López-Sánchez MJ, Neef A, Peretó J, Patiño-Navarrete R, Pignatelli M, Latorre A, Moya A: **Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*.** *PLoS Genet* 2009, **5**:e1000721.
- Sabree ZL, Kambhampati S, Moran NA: **Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont.** *Proc Natl Acad Sci USA* 2009, **106**:19521-1956.
- Gruwell ME, Morse GE, Normark BB: **Phylogenetic congruence of armored scale insects (Hemiptera: Diaspididae) and their primary endosymbionts from the phylum Bacteroidetes.** *Mol Phylogenet Evol* 2007, **44**:267-280.
- Gottlieb Y, Ghanim M, Gueguen G, Kontsedalov S, Vavre F, Fleury F, Zchori-Fein E: **Inherited intracellular ecosystem: symbiotic bacteria share bacteriocytes in whiteflies.** *FASEB J* 2008, **22**:2591-2599.
- Stingl U, Maass A, Radek R, Brune A: **Symbionts of the gut flagellate *Staurojoenina* sp. from *Neotermes cubanus* represent a novel, termite-associated lineage of Bacteroidales: description of '*Candidatus Vestibaculum illigatum*'.** *Microbiology* 2004, **150**:2229-2235.
- Sabree ZL, Degnan PH, Moran NA: **Chromosome stability and gene loss in cockroach endosymbionts.** *Appl Environ Microbiol* 2010, **76**:4076-4079.
- Grimaldi D, Engel MS: **Evolution of Insects.** New York/Cambridge: Cambridge University Press; Grimaldi D, Engel MS 2005.
- Cochran DG: **Nitrogen excretion in cockroaches.** *Annu Rev Entomol* 1985, **30**:29-49.
- Mullins DE, Cochran DG: **Nitrogen excretion in cockroaches: uric acid is not a major product.** *Science* 1972, **177**:699-701.
- Mullins DE, Cochran DG: **Nitrogen metabolism in the American cockroach: an examination of whole body and fat body regulation of cations in response to nitrogen balance.** *J Exp Biol* 1974, **61**:557-570.
- O'Donnell M: **Insect excretory mechanisms.** In *Advances in Insect Physiology. Volume 35.* New York: Academic Press; Simpson SJ 2008:1-122.
- Needham J: **Contributions of chemical physiology to the problem of reversibility in evolution.** *Biol Rev* 1938, **13**:225-251.
- Cochran DG, Mullins DE, Mullins KJ: **Cytological changes in the fat body of the American cockroach, *Periplaneta americana*, in relation to dietary nitrogen levels.** *Ann Entomol Soci Amer* 1979, **72**:197-205.
- Moya A, Peretó J, Gil R, Latorre A: **Learning how to live together: genomic insights into prokaryote-animal symbioses.** *Nat Rev Genet* 2008, **9**:218-229.
- Moran NA, McCutcheon JP, Nakabachi A: **Genomics and evolution of heritable bacterial symbionts.** *Annu Rev Genet* 2008, **42**:165-190.
- Lamelas A, Gosálbes MJ, Moya A, Latorre A: **New clues about the evolutionary history of metabolic losses in bacterial endosymbionts, provided by the genome of *Buchnera aphidicola* from the aphid *Cinara tujaefilina*.** *Appl Environ Microbiol* 2011, **77**:4446-4454.
- Edwards JS, Covert M, Palsson B: **Metabolic modelling of microbes: the flux-balance approach.** *Environ Microbiol* 2002, **4**:133-140.
- Covert MW, Palsson BO: **Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*.** *J Biol Chem* 2002, **277**:28058-28064.
- Puchalka J, Oberhardt MA, Godinho M, Bielecka A, Regenerhardt D, Timmis KN, Papin JA, Martins dos Santos V: **Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology.** *PLoS Comput Biol* 2008, **4**:e1000210.
- Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, Palsson BO: **Genome-scale metabolic model of *Helicobacter pylori* 26695.** *J Bacteriol* 2002, **184**:4582-4593.
- Orth JD, Thiele I, Palsson BØ: **What is flux balance analysis?** *Nat Biotechnol* 2010, **28**:245-248.
- Thiele I, Palsson BØ: **A protocol for generating a high-quality genome-scale metabolic reconstruction.** *Nat Protoc* 2010, **5**:93-121.
- Locke M: **The fat body.** In *Microscopic anatomy of invertebrates. Insecta Mundi. Volume 11B.* New York: Wiley/Harrison FW, Locke M 1998:641-686.
- Thomas GH, Zucker J, Macdonald SJ, Sorokin A, Goryanin I, Douglas AE: **A fragile metabolic network adapted for cooperation in the symbiotic bacterium *Buchnera aphidicola*.** *BMC Syst Biol* 2009, **3**:24.
- Pál C, Papp B, Lercher MJ, Csérmely P, Oliver SG, Hurst LD: **Chance and necessity in the evolution of minimal metabolic networks.** *Nature* 2006, **440**:667-670.
- Yizhak K, Tuller T, Papp B, Ruppin E: **Metabolic modeling of endosymbiont genome reduction on a temporal scale.** *Mol Syst Biol* 2011, **7**:479.
- Ates O, Toksoy Oner E, Arga KY: **Genome-scale reconstruction of metabolic network for a halophilic extremophile, *Chromohalobacter salexigens* DSM 3043.** *BMC Syst Biol* 2011, **5**:12.
- Oberhardt MA, Puchalka J, Fryer KE, Martins dos Santos VA, Papin JA: **Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1.** *J Bacteriol* 2008, **190**:2790-2803.
- Zhang Y, Thiele I, Weekes D, Li Z, Jaroszewski L, Ginalski K, Deacon AM, Woolley J, Lesley SA, Wilson IA, Palsson B, Osterman A, Godzik A: **Three-dimensional structural view of the central metabolic network of *Thermotoga maritima*.** *Science* 2009, **325**:1544-1549.
- Kiers ET, Rousseau RA, West SA, Denison RF: **Host sanctions and the legume-rhizobium mutualism.** *Nature* 2003, **425**:78-81.
- Burgard AP, Nikolaev EV, Schilling CH, Maranas CD: **Flux coupling analysis of genome-scale metabolic network reconstructions.** *Genome Res* 2004, **14**:301-312.
- Suthers PF, Zomorodi A, Maranas CD: **Genome-scale gene/reaction essentiality and synthetic lethality analysis.** *Mol Syst Biol* 2009, **5**:301.
- Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO: **A genome-scale metabolic**

- reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 2007, **3**:121.
34. Gil R, Silva FJ, Peretó J, Moya A: **Determination of the core of a minimal bacterial gene set.** *Microbiol Mol Biol Rev* 2004, **68**:518-537.
  35. Gabaldon T, Peretó J, Montero F, Gil R, Latorre A, Moya A: **Structural analyses of a hypothetical minimal metabolism.** *Philos Trans R Soc Lond B Biol Sci* 2007, **362**:1751-1762.
  36. Williams LE, Wernegreen JJ: **Unprecedented loss of ammonia assimilation capability in a urease-encoding bacterial mutualist.** *BMC Genomics* 2010, **11**:687.
  37. Reed JL, Vo TD, Schilling CH, Palsson BO: **An expanded genome-scale model of *Escherichia coli* K-12 (IJR904 GSM/GPR).** *Genome Biol* 2003, **4**:R54.
  38. Chen F, Mackey AJ, Stoecckert CJ Jr, Roos DS: **OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups.** *Nucleic Acids Res* 2006, **34**:363-368.
  39. Bernardet JF, Nakagawa Y: **An introduction to the family *Flavobacteriaceae*.** In *The Prokaryotes: a Handbook on the Biology of Bacteria. Volume 7.* 3 edition. New York: Springer-Verlag; Dworkin M et al 2006:455-480.
  40. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
  41. Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: Identification of signaling domains.** *Proc Natl Acad Sci USA* 1998, **95**:5857-5864.
  42. Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP, Bork P: **Recent improvements to the SMART domain based sequence annotation resource.** *Nucleic Acids Res* 2002, **30**:242-244.
  43. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.** *Nat Protoc* 2007, **2**:727-738.
  44. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J, SBML Forum: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.** *Bioinformatics* 2003, **19**:524-531.

doi:10.1186/1471-2180-12-S1-S5

**Cite this article as:** González-Domenech et al: Metabolic stasis in an ancient symbiosis: genome-scale metabolic networks from two *Blattabacterium cuenoti* strains, primary endosymbionts of cockroaches. *BMC Microbiology* 2012 **12**(Suppl 1):S5.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

