FEBS 24885 FEBS Letters 498 (2001) 135–139

Minireview

Mutualists and parasites: how to paint yourself into a (metabolic) corner

Ivica Tamas, Lisa M. Klasson, Jonas P. Sandström, Siv G.E. Andersson*

Department of Molecular Evolution, University of Uppsala, Norbyvägen 18C, S-752 36 Uppsala, Sweden

Received 14 April 2001; accepted 27 April 2001

First published online 10 May 2001

Edited by Gunnar von Heijne

Abstract Eukaryotes have developed an elaborate series of interactions with bacteria that enter their bodies and/or cells. Genome evolution of symbiotic and parasitic bacteria multiplying inside eukaryotic cells results in both convergent and divergent changes. The genome sequences of the symbiotic bacteria of aphids, *Buchnera aphidicola*, and the parasitic bacteria of body louse and humans, *Rickettsia prowazekii*, provide insights into these processes. Convergent genome characteristics include reduction in genome sizes and lowered G+C content values. Divergent evolution was recorded for amino acid and cell wall biosynthetic genes. The presence of pseudogenes in both genomes provides examples of recent gene inactivation events and offers clues to the process of genome deterioration and host-cell adaptation.

Key words: Genome evolution; Degeneration; Intracellular parasite; Intracellular symbiont; *Rickettsia*; *Buchnera*

1. Introduction

Organelles represent the ultimate example of bacterial domestication by eukaryotic cells [1]. Parasitic and symbiotic bacteria that multiply inside eukaryotic cells represent earlier stages of a similar integration process. The distinction between parasites and symbionts is mainly defined by the effect the bacterium has on its eukaryotic host. From the bacterial point of view, however, both lifestyles give rise to similar problems that have to be solved by homologous or analogous systems. For example, both organisms have to adhere to and enter their host cells; they have to avoid or counteract the hosts' defence systems; they have to multiply; and they have to exit their host cells to colonise new cells and/or new hosts. The fine-tuning of this interaction results in differences that are characteristic of the two lifestyles. As complete bacterial genome sequence data are now available from several obligate intracellular bacteria, it is possible to compare these differences from a genomic perspective.

Buchnera belongs to the γ subdivision of the Proteobacteria [2]. These bacteria are obligate endosymbionts residing within specialised cells termed bacteriocytes of their aphid hosts [3] (Fig. 1A). Bacteriocytes are arranged in clusters of cells forming a bilobed structure within an aphid body cavity known as the bacteriome [4]. On average an aphid adult contains around 1.0 million *Buchnera* cells per mg aphid [5]. Infection of the host is maintained in a strictly vertical manner via maternal passage of the *Buchnera* cells to eggs and embryos [6]. *Buchnera* has a monophyletic origin, estimated to be at least 150–250 million years old [7].

Rickettsia belongs to the α subdivision of the Proteobacteria [8,9]. Members of this genus have an obligate intracellular lifestyle and some but not all are pathogenic to humans. They infect vertebrate hosts with the help of bloodsucking parasitic arthropods such as fleas, lice and ticks. The parasite *R. prowazekii*, the causative agent of epidemic typhus, is transmitted to humans via the human body louse (Fig. 1B), whereas endemic typhus, caused by *R. typhi*, is spread to humans by rat fleas [10]. Members of the spotted fever group rickettsia are transmitted to humans from their animal hosts by ticks [10]. Some species multiply exclusively in the host cell cytoplasm whereas others can also grow in the cell nucleus [11].

In this paper, we compare the symbiont *Buchnera* and the parasite *Rickettsia* with special emphasis on their metabolic profiles, as inferred from the genome sequences.

2. The cost of metabolic exchange is paid in nucleotides

Rickettsia and *Buchnera* have many features in common, the most characteristic of which is their drastically reduced genome sizes. The complete genome size of *R. prowazekii* is 1.1 Mb [12] and that of *Buchnera* sp. is even smaller, only 0.64 Mb [13]. In total, 583 and 834 genes have been identified in the genomes of *Buchnera* and *Rickettsia*, respectively. These genomes are thought to be the modern remnants of a long-term reduction process from much larger genomes of their free-living ancestors. Indeed, the metabolic profile of *Buchnera* seems to be nothing but a 14% subset of the metabolic capacity encoded by the 4.7 Mb genome of *Escherichia coli* [14] (Fig. 2). Only four genes have been identified that are uniquely present in *Buchnera*, subsp. *Acyrthosiphon pisum* (Ap) [13]. Thus, the cost for protection and supply of nutrients is paid directly in nucleotides – irrespective of whether the relationship is based on a parasitic or a symbiotic relationship.

3. Metabolic integration versus exploitation

Even though *Buchnera* and *Rickettsia* have learned how to reproduce in almost identical environments, they interact with their hosts in strikingly different ways. While the parasitic bacterium forces the host cell to feed it with metabolic products, the symbiotic bacterium will rather have the same com-

*Corresponding author. Fax: (46)-18-471 64 04.
E-mail: siv.andersson@ebc.uu.se

PII: S0014-5793(01)02459-0

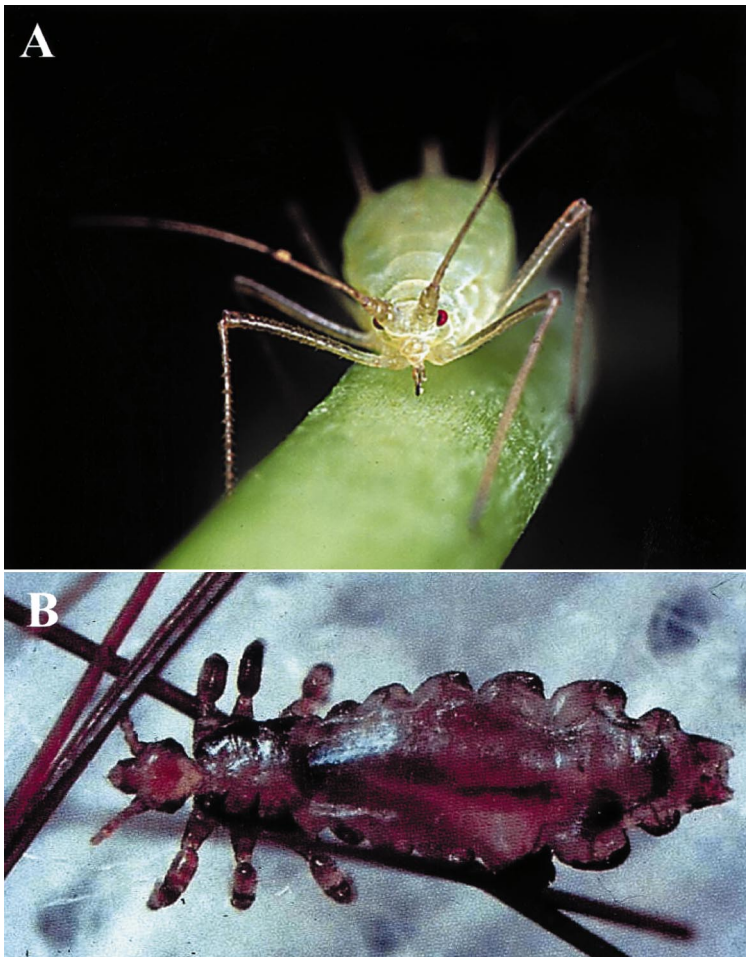

Fig. 1. A: The pea aphid (*Acyrthosiphon pisum*), one host of the mutualistic bacterium *Buchnera* sp. B: The human louse (*Pediculus humanus*), host and vector of the bacterium *Rickettsia prowazekii*.

pounds given in exchange for something else. Indeed, the main difference between *Rickettsia* and *Buchnera* is found in their capabilities to produce metabolic molecules, transporting these compounds and building cell envelope structures (Fig. 3). The symbiotic relationship is reflected in a higher fraction of genes allocated to biosynthetic functions, whereas the parasitic relationship is characterised by a higher fraction of transport functions. Below, we discuss these differences in more detail.

3.1. Mutual exchange of amino acids

The symbiotic relationship between *Buchnera* and its aphid host is mostly built on the symbiont's ability to produce amino acids for the host. Not surprisingly, the genome of *Buchnera* contains genes encoding host-essential amino acids, whereas those for non-essential amino acids are missing [13]. Genes involved in the biosynthesis of leucine and tryptophan have been placed on plasmids, so that the expression levels can easily be increased by copy number effects [4]. The retention and amplification of the essential amino acid genes in *Buchnera* is probably due to strong positive host selection.

This is in striking contrast to parasitic bacteria like *Rickettsia* that predates on the host production of amino acids, and consequently, only a few amino acid biosynthetic genes have been identified in the *Rickettsia* genome [12].

3.2. Import and export of metabolic compounds

To compensate for the lack of amino acid biosynthetic genes, the *R. prowazekii* genome contains as many as 15 genes encoding amino acid transporters [12]. This seems to be part of a general strategy since *R. prowazekii* has a much broader spectrum of transport proteins than *Buchnera*. What is remarkable about this is perhaps not that there are so many genes encoding transporters in *Rickettsia*, but that there are so few of these genes in *Buchnera*. The genome sequence shows that *Buchnera* cannot synthesise all compounds required for growth, implying that many metabolic products and possibly even enzymes have to be imported from the host cell cytoplasm. To support the symbiotic relationship there is also a requirement for a system to export compounds essential for host survival, for example amino acids. Thus, transport systems must be present in *Buchnera*, but what proteins are responsible for the import and export functions?

The *Buchnera* genome contains a surprisingly large number of genes encoding flagellar proteins [13]. These have previously been suggested to be part of a possible transportation apparatus in *Salmonella typhimurium* [15] and in *Yersinia enterocolitica*. In the latter the flagellum was shown to function both in the uptake and in the secretion of proteins [16]. The flagellar genes in *Buchnera* are situated in two large operons, *flg* and *fli*, which consist of 12 genes each. In addition, there is

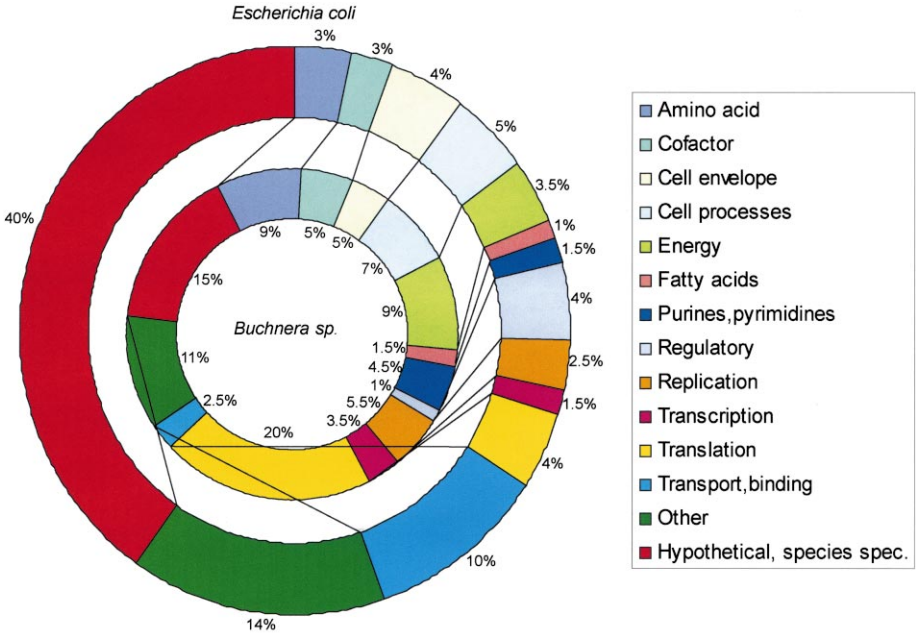

Fig. 2. Comparison of the genome content from *E. coli* [15] and *Buchnera* sp. [13]. Genes are divided into functional categories and presented as relative fractions of the whole genome.

a short operon containing two flagellar genes, *flhA* and *flhB*. Since the 24 flagellar genes occupy a relatively large fraction of the genome they must serve an important functional role, possibly related to transport functions. The alternative interpretation is that these transport functions are supplied by the host organism.

3.3. Envelope structures: two alternative routes towards specialisation

In addition to communicating with their hosts by supplying or stealing small metabolites, both parasites and mutualists have evolved systems for promoting (or avoiding) direct, physical interactions with their host cells. Symbiotic bacteria like *Buchnera* are safely embedded in both a host-derived membrane and its own membrane [4]. This host cell protection is reflected in an unusually small number of genes encoding cell membrane components and fatty acids in *Buchnera* [13]. Only five genes (*smpA, lpcA, ompA, ompF* and *rfaE*) encode proteins directly associated with the cell membrane biosynthesis or enzymes that produce molecules attached to the cell membrane.

In contrast, *Rickettsia* and other intracellular parasites have evolved elaborate mechanisms for changing their cell surface structures [12], so as to avoid being recognised and killed by the host. The lack of the corresponding set of cell surface proteins makes the symbiont extremely vulnerable – if the host ever turned against it, *Buchnera* would not have a chance to survive.

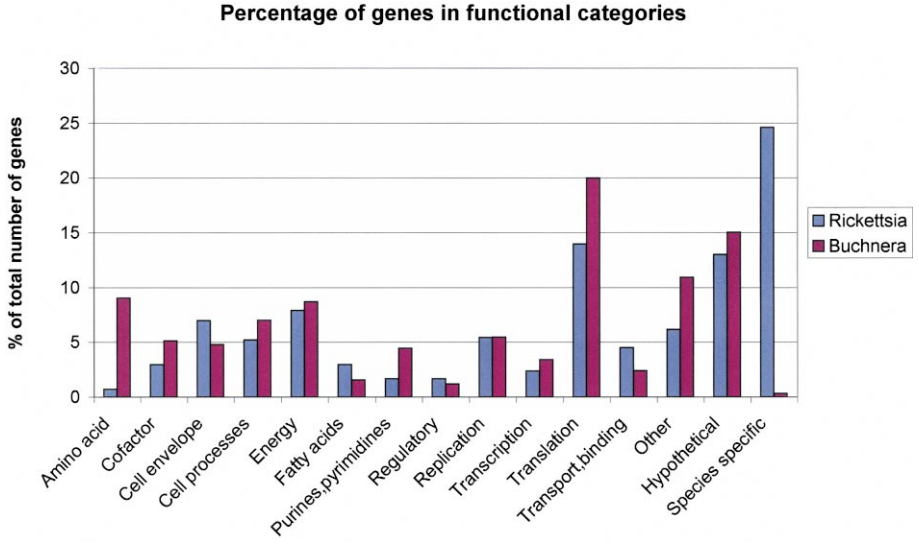

Fig. 3. Comparison of the genome content from *R. prowazekii* [12] and *Buchnera* sp. [13]. Genes are divided into functional categories and presented as relative fractions of the whole genome. The last category refers to hypothetical genes that are specific for each species.

4. Bioenergetic profiles

Even though *Rickettsia* and *Buchnera* are both capable of ATP production, their systems for energy production are complementary rather than identical. *Buchnera* has a full set of genes for glycolysis with the exception of hexokinase, which phosphorylates glucose to glucose 6-phosphate [13]. The glucose may be directly imported from the host cell and phosphorylated at the same time. Indeed, two genes for glucose-specific uptake, *crr* and *ptsG*, have been detected in the genome. In contrast, the glycolytic pathway is completely missing in *R. prowazekii* [12] while the trichloroacetic acid (TCA) cycle, missing in *Buchnera*, is intact. Only two genes from the TCA cycle have been detected in the genome of *Buchnera*, *sucA* and *sucB* [13].

Genes for the pyruvate dehydrogenase complex are present in both organisms suggesting that pyruvate, produced by glycolysis in *Buchnera* and imported from the cytosol by *Rickettsia*, can be converted into acetyl-CoA in both organisms. However, the gene sets are strikingly different; while *Buchnera* has the same gene set as *E. coli* (*aceE, aceF* and *lpdA*) [13,14], *R. prowazekii* uses a gene set similar to that of mitochondria and Gram-positive bacteria [12]. Both genomes have full sets of genes for the NADH dehydrogenase complex and for the ATP synthase complex. *R. prowazekii* contains several genes encoding cytochrome bc_1 reductase complex, whereas *Buchnera* only encodes genes for cytochrome O ubiquinol oxidase which in *E. coli* predominates when cell are grown at high aeration.

So, although *Buchnera* and *Rickettsia* use somewhat different systems for ATP production, both seem to be self-sufficient in terms of energy production. However, a striking difference is that *R. prowazekii* has also learned how to exploit cytosolic ATP with the help of ATP/ADP transporters [12], while *Buchnera* does not appear to have evolved mechanisms for exploiting the host cell supply of energy.

5. Specialisation and genome degradation: to be or not to be an organism

Whereas the complete loss of a function can only be indirectly inferred from the current gene repertoires, pseudogenes can be used to follow the deterioration process in real time. We have previously identified 12 pseudogenes in *R. prowazekii*, all of which show signs of degradation in at least one other *Rickettsia* species [12,17,18]. Likewise, eight pseudogenes have been described in *Buchnera* (Ap) [13], and an even larger number of pseudogenes have been identified in *Buchnera*, subsp. *Schizaphis graminium* (Sg) (Tamas, Klasson, Sandström, Moran and Andersson, unpublished).

The pseudogenes fall into several different categories. Some contain just one or a few irregularities such as small insertions and deletions, substitutions that convert codons to termination codons as well as the occasional loss of start and/or stop codons. Others are much more extensively degraded and are recognised as pseudogenes only because of sequence similarities to a gene or a less degraded pseudogene in a closely related species within the genera *Buchnera* and *Rickettsia*.

Many genes seem to have been inactivated as a consequence of a deepened interaction between the bacterium and the host. For example, a biosynthetic gene, *metK*, which encodes an enzyme involved in the production of the metabolite *S*-adenosylmethionine (SAM), has been inactivated in *Rickettsia* [12,18]. It seems likely that *Rickettsia* has learned how to import SAM from the host cell cytoplasm, thereby making the corresponding biosynthetic gene unnecessary. Likewise, a set of genes involved in the biosynthesis of amino acids has recently been inactivated in *Buchnera* (Sg) (Tamas, Klasson, Sandström, Moran and Andersson, unpublished).

Other pseudogenes in *Buchnera* include genes involved in cell wall biosynthesis. For example, several genes involved in murein biosynthesis contain mutations in *Buchnera* (Sg) (Tamas, Klasson, Sandström, Moran and Andersson, unpublished), in contrast to the *R. prowazekii* genome with a fully preserved set of murein biosynthetic genes [12]. Other genes involved in cell wall biosynthesis and maintenance have also been identified as pseudogenes in *Buchnera* (Sg) (Tamas, Klasson, Sandström, Moran and Andersson, unpublished). The lack of genes involved in the biosynthesis of phospholipids – which are otherwise indispensable building blocks of the membrane – provides additional evidence to suggest that *Buchnera* has an impaired capacity to build cell envelope structures.

Taken together, this indicates that *Buchnera* may be in the process of completely abandoning cell envelope biosynthesis. Yet another bacterial structure may be in the process of being replaced by host-derived components, similar to the many eukaryotic proteins that have replaced the ancestral bacterial proteins in the mitochondrion [1,19]. As more and more genomes from invading and engulfed bacteria are now being sequenced, the boundary between an organism and an integral cellular component becomes increasingly diffuse and difficult to define.

6. Perspective

The gradual process of host cell adaptation and integration may eventually push the *Buchnera* genomes towards the small size of the mitochondrial and chloroplast genomes. Vice versa, one may view the current *Rickettsia*–louse and *Buchnera*–aphid associations as examples of what the mitochondrial and chloroplast genomes may once have looked like. Many of the genetic losses are probably irreversible and trigger a pathway of no return for the invading bacterium. However, the genetic changes awaiting a would-be intracellular bacterium arise from the nature of interaction with its host and its environment. Whether the bacterium is initially recognised by the host as being either a parasite or a commensal mutualist will most likely play a crucial role in its future evolution. The many facultative intracellular parasites are examples of one of these exploratory pathways, in which virulence genes may be gained from other organisms to facilitate the process of uptake and internalisation.

At some stage the association becomes obligate, after which a massive loss of genes is initiated. The obligate intracellular parasite *Mycobacterium leprae*, which has a genome size of 3.2 Mb, a coding content of only 50% and as many as 1116 pseudogenes [20], provides a wonderful example of the stage at which rapid deterioration occurs. Eventually, the rate of degradation will slow down and the genome size will be reduced at a slower and slower pace. The small genomes of *Rickettsia* and *Buchnera* are apparently still in the process of degradation, as evidenced by the presence of pseudogenes and non-coding DNA in both genomes, although the

fraction of deteriorating genes is much smaller than in *M. leprae*.

Once genes that are potentially valuable for the establishment of a particular association are lost, there is no viable means for their return in a closed intracellular environment. The final step of this minimisation process could therefore in theory lead to the loss of all genes, except those that are essential for keeping the host–bacterial conglomerate reproducing. From now on, the intracellular bacterium will be locked in its metabolic corner and therefore extremely vulnerable to any environmental changes. If lice and aphids are eliminated from the Earth, obligate intracellular parasites and symbionts such as *Rickettsia* and *Buchnera* will immediately be driven to extinction. On the other hand, if the host–bacterial conglomerate is very successful, the bacterium may survive within its closed metabolic corner for hundreds of millions of years – and eventually become so well integrated with its host that it is no longer recognised as a separate entity.

Acknowledgements: The picture of the human louse (*Pediculus humanus*) was a kind gift from A.-M. Rasmussen, Danish Pest Infestation Laboratory (DPIL). The authors' work is supported by the Swedish Council for Forestry and Agricultural Research, the Foundation for Strategic Research, Sweden and the National Science Foundation, USA.

References

[1] Kurland, C.G. and Andersson, S.G.E. (2000) Mol. Microbiol. Rev. 64, 786–820.
[2] Munson, M.A., Baumann, P. and Kinsey, M.G. (1991) Int. J. Syst. Bacteriol. 41, 566–568.
[3] Buchner, P. (1965) Endosymbiosis of Animals with Plant Microorganisms, pp. 210–332, Interscience, New York.
[4] Baumann, P., Baumann, L., Chi-Yung, L., Rouhbakhsh, D., Moran, N.A. and Clark, M. (1995) Annu. Rev. Microbiol. 49, 55–94.
[5] Douglas, A.E. and Prosser, W.A. (1992) J. Insect Physiol. 38, 565–568.
[6] Hinde, R. (1971) J. Invertebr. Pathol. 17, 333–338.
[7] Moran, N.A., Munson, M.A., Baumann, P. and Ishikawa, H.A. (1993) Proc. R. Soc. Lond. B 253, 167–171.
[8] Weisburg, W.G., Dobson, M.E., Samuel, J.E., Dasch, G.A., Mallavia, L.P., Baca, O., Mandelco, L., Sechrestr, J.E. and Weiss, E. (1989) J. Bacteriol. 171, 4202–4206.
[9] Olsen, G.J., Woese, C.R. and Overbeek, R. (1994) J. Bacteriol. 176, 1–6.
[10] Raoult, D. and Roux, V. (1997) Clin. Microbiol. Rev. 10, 694–719.
[11] Wisseman, C.L.J., Edlinger, E.A., Waddell, A.D. and Jones, M.R. (1976) Infect. Immun. 14, 1052–1064.
[12] Andersson, S.G.E., Zomorodipour, A., Andersson, J., Siheritz-Ponten, T., Alsmark, C.M., Podowski, R.M., Näslund, K.A., Eriksson, A., Winkler, H.H. and Kurland, C.G. (1998) Nature 396, 133–140.
[13] Shigenobu, S., Watanabe, H., Hattori, M., Sasaki, Y. and Ishikawa, I. (2000) Nature 407, 81–86.
[14] Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J.E. and Aizawa, S. (1998) Science 280, 602–605.
[15] Blattner, F.R., Plunkett III, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997) Science 277, 1453–1474.
[16] Young, G.M., Schmiel, D.H. and Miller, V.L. (1999) Proc. Natl. Acad. Sci. USA 96, 6456–6461.
[17] Andersson, J.O. and Andersson, S.G.E. (1999) Mol. Biol. Evol. 16, 1178–1191.
[18] Andersson, J.O. and Andersson, S.G.E. (2000) Mol. Biol. Evol. 18, 825–837.
[19] Karlberg, O., Canbäck, B., Kurland, C.G. and Andersson, S.G.E. (2000) Yeast 17, 170–187.
[20] Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J.R. and Barrell, B.G. (2001) Nature 409, 1007–1011.