

FEMS Microbiology Reviews 29 (2005) 99-117



www.fems-microbiology.org

Some lessons from Rickettsia genomics

Patricia Renesto^a, Hiroyuki Ogata^b, Stéphane Audic^b, Jean-Michel Claverie^b, Didier Raoult^{a,*}

^a Unité des Rickettsies, CNRS UMR 6020, IFR-48, Faculté de Médecine – Universite de la Mediterranee, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France

^b Information Génomique et Structurale, UPR 2589, 31 Chemin Joseph Aiguier, 13402 Marseille, France

Received 28 May 2004; received in revised form 11 August 2004; accepted 2 September 2004

First published online 29 September 2004

Abstract

Sequencing of the *Rickettsia conorii* genome and its comparison with its closest sequenced pathogenic relative, i.e., *Rickettsia prowazekii*, provided powerful insights into the evolution of these microbial pathogens. However, advances in our knowledge of rickettsial diseases are still hindered by the difficulty of working with strict intracellular bacteria and their hosts. Information gained from comparing the genomes of closely related organisms will shed new light on proteins susceptible to be targeted in specific diagnostic assays, by new antimicrobial drugs, and that could be employed in the generation of future rickettsial vaccines. In this review we present a detailed comparison of the metabolic pathways of these bacteria as well as the polymorphisms of their membrane proteins, transporters and putative virulence factors. Environmental adaptation of *Rickettsia* is also discussed. © 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Rickettsia; Comparative genomics; Metabolic pathways

Contents

1.	Introduction	100
2.	Methods	101
3.	Major metabolite pathway analysis	105
	3.1. Carbohydrate metabolism	105
	3.2. Energy metabolism	105
	3.3. Lipid metabolism	106
	3.4. Nucleotide metabolism	106
	3.5. Amino acid metabolism	107
	3.6. Glutathione metabolism	108
	3.7. Metabolism of complex carbohydrates	108
	3.8. Metabolism of complex lipids	109
	3.9. Metabolism of cofactors and vitamins.	109
4.	Rickettsia membranes.	109
5.	Transporters	111

* Corresponding author. Tel.: +33 4 91 38 55 17;

fax: +33 4 91 38 77 72.

E-mail address: didier.raoult@medecine.univ-mrs.fr (D. Raoult).

0168-6445/\$22.00 © 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved. doi:10.1016/j.femsre.2004.09.002

	.1. ABC transporters	. 111
	2. Multidrug transporters	111
	.3. Protein and peptide secretion	111
	.4. Amino acid transporters	112
	5.5. Nucleoside and nucleotide transporters	112
6.	Virulence factors	112
7.	Environmental adaptation of <i>Rickettsia</i>	113
	1. DNA repair mechanisms	113
	2.2. Genomic evolution	113
8.	Concluding remarks	114
	Acknowledgements	114
	References	115

1. Introduction

Rickettsia species are small, Gram-negative bacilli, consisting of two antigenically defined groups: the spotted fever group and the typhus group [1]. These bacteria are phenotypically characterized by their strict intracellular location and their association with arthropods [2]. Despite numerous efforts [3–5] genetic manipulation of Rickettsia remained impossible for years [6]. In a work published few months ago, Qin et al. [7] isolated for the first time stable Rickettsia prowazekii insertional knock-out mutants. Their approach was based on the use of the transposome system which efficiently generated random chromosomal insertions. Selection for rickettsial transformants was provided by the arr-2 gene which encodes for an enzyme that ADP-ribosylates rifampin, thereby destroying its antibacterial activity, and which was inserted in the transposon under the control of the R. prowazekii rpsL promoter. This method provides the first efficient pathway for identifying nonessential genes of Rickettsia. Until this achievement, the absence of convenient tools has hampered the unequivocal identification of the genes involved in the pathogenicity of these bacteria for which no efficient vaccine is actually available. Recent outbreaks [8,9] testify to the real threat that R. prowazekii poses. Furthermore, people may carry rickettsiae asymptomatically after the initial infection and can develop Brill-Zinsser disease which is the relapsing bacteremic form of the infection [10]. The recent inclusion of rickettsiae in the list of potential bioterrorism agents [11] has made the development of tools to prevent, diagnose and treat these infections a new priority.

During the last few years, two rickettsial genomes have been sequenced, i.e., those of *R. prowazekii* [12] and *Rickettsia conorii* [13], which are responsible for typhus and Mediterranean spotted fever, respectively [2]. These genomic sequences provided significant insights into the molecular physiology and evolution of these bacteria. In the course of evolution, their tight associations with host physiology and ecology caused the elimination of biosynthetic metabolic pathways for superfluous compounds or compounds deemed beneficial but non-essential. A number of these synthetic pathways have been replaced by transport systems. As a consequence of this ongoing evolutionary process [14], these bacteria possess small genomes. Comparison of the rickettsial genome sequences precisely characterized the close phylogenetic relationship between the modern mitochondria and rickettsiae and demonstrated a similarity in their functional profiles such as in the ATP production pathway [12]. As in other intracellular microorganisms, few lateral gene transfer events were evident [13,15,16]. A low rate of genetic exchange by obligate intracellular bacteria is probably due to their isolated location within their host cells [17], and might limit the mode of genome evolution compared to other free living bacteria. Exceptional cases of possible lateral gene acquisitions include ATP/ADP translocases that probably entered rickettsial genomes by horizontal transfer from Chlamydia [18]. This acquisition probably facilitated the survival of the bacterium within the host. A comparative analysis of the two rickettsial genomes revealed that R. prowazekii is a quasi subset of R. conorii, as illustrated in the Fig. 1. Homologues for 540 of the 1374 R. conorii ORFs (39%) were not found within the R. prowazekii genome. The order of orthologous genes along the chromosome is highly conserved between R. conorii and R. prowazekii with the exception of a handful of genome rearrangements near the replication terminus. This may result from the absence of *recB*, *recC* and recD genes, responsible for chromosomal inversions in Escherichia coli [19]. Comparative genome analyses also revealed similarities and differences in functional complements between R. conorii and R. prowazekii. While shared core genes mostly encode metabolic and cellular processes, strain-specific genes include genes for cell surface proteins or virulence factors (Table 1). The variations in the genomic repertoires of rickettsiae might be associated with their capacity to survive under different circumstances during their long-term infection of genetically diverse hosts such as ticks (R. conorii) or



Fig. 1. Comparative analysis of *R. prowazekii* and *R. conorii* genomes. Synthetic display of protein coding gene differences between *R. conorii* and *R. prowazekii*: Only genes not found in the cognate genome are displayed. Genes are displayed as filled boxes with the following colors: green, gene of unknown function; red, gene of known or putative function, and purple, pseudo-genes and split genes. The region of genomic rearrangement between the two species is displayed as a grey shaded area.

lice (*R. prowazekii*). Genomic sequence information, particularly the polymorphisms of encoded membrane proteins should be exploited to provide rapid and accurate diagnostics to identify such pathogens.

Elucidation of nutritional requirements of *Rickettsia* in vivo as well as the mechanisms by which these bacteria acquire these nutrients inside a host is critical in understanding virulence and disease. For example, if a pathogen is unable to obtain the precursors for the nucleic or amino acids required for growth, it must stop proliferating. In this respect, metabolite pathways – or transport mechanisms – can be considered as potential therapeutic targets of rickettsial diseases. In this review we propose a detailed reconstruction of the metabolic capabilities and transporter profiles of *Rickettsia*. Genes encoding the most commonly recognized virulence determinants such as membrane proteins, or toxins are examined. Finally, physiology and adaptability of these bacteria to various environmental conditions are also detailed.

2. Methods

The metabolic and gene content analyses of *Rickett-sia* presented in this review were based on the Kyoto encyclopedia of genes and genomes (KEGG) metabolic database [20] and on works cited in the text. The KEGG metabolic map identifiers were indicated in the text. The gene function classification employed was also derived from the KEGG web site. Additional information about transporters was obtained from the TransportDB [21].

Table 1

List of genes differentially identified within R. conorii (Rco) or R. prowazekii (Rpr) genomes

- 1 Replication
- DNA replication, modification, recombination and repair (50 Rco/ 41 Rpr) 1.1 *rco:RC0870 dat; methylated-DNA-protein-cysteine methyltransferase [EC:2.1.1.63] {DEGRADATION} *rco:RC1038 mutM; formamidopyrimidine-DNA glycosidase [EC:3.2.2.23] *rco:RC0525 mutT; similar to mutator protein MutT [EC:3.6.1.-] *rco:RC0690 unknown (PARTIAL) similarity to DNA adenine methyltransferases (MboA) {DEGRADATION} *rco:RC1278 unknown (PARTIAL) split gene of DNA repair protein (RadC. {SPLIT: RC1277 + RC1278 = sp:RADC_RHOCA} *rco:RC1277 unknown (PARTIAL) split gene of DNA repair protein (RadC. {SPLIT: RC1277 + RC1278 = sp:RADC_RHOCA} {DEGRADATION} *rco:RC0868 unknown (PARTIAL) split gene of methylated-DNA-protein-cysteine methyltransferase {SPLIT: RC0867 + RC0868 + RC0869 = pae:PA3596} *rco:RC0869 unknown (PARTIAL) split gene of methylated-DNA-protein-cysteine methyltransferase {SPLIT: RC0867 + RC0868 + RC0869 = pae:PA3596} *rco:RC0867 unknown (PARTIAL) split gene of methylated-DNA-protein-cysteine methyltransferase {SPLIT: RC0867 + RC0868 + RC0869 = pae:PA3596} Degradation of DNA (4 Rco/4 Rpr) 1.2 1.3 Other (7 Rco/0 Rpr) *rco:RC1344 dinJ; DNA-damage-inducible protein J *rco:RC0490 phage-related protein {DEGRADATION} *rco:RC0688 putative transposable insertion element {DEGRADATION} *rco:RC1318 similarity to plasmid stability protein *rco:RC1350 similarity to type I restriction enzyme M subunit *rco:RC1348 similarity to type I restriction enzyme S subunit *rco:RC0873 unknown (PARTIAL) similarity to type I restriction enzyme ristriction subunit {VERY_UNIQUE} 2. Transcription 2.1 RNA synthesis and modification (13 Rco/13 Rpr) 2.2 Degradation of RNA (9 Rco/9 Rpr) 3 Translation 3.1 Aminoacyl-tRNA synthetases (21 Rco/21 Rpr) 3.2 tRNA and amino acyl-tRNA modification (19 Rco/17 Rpr) *rco:RC0674 def2; polypeptide deformylase [EC:3.5.1.31] *rco:RC0080 def3; polypeptide deformylase [EC:3.5.1.31] *rco:RC0160 tRNA/rRNA methyltransferase [EC:2.1.1.-] {DEGRADATION} 3.3 Ribosomal proteins, synthesis and modification (57 Rco/57 Rpr) Protein modification and translation factors (16 Rco/14 Rpr) 3.4 *rco:RC1367 unknown (PARTIAL) split gene of O-linked GlcNAc transferase {SPLIT: RC1366 + RC1367 = mth:MTH83} *rco:RC1366 unknown (PARTIAL) split gene of O-linked GlcNAc transferase {SPLIT: RC1366 + RC1367 = mth:MTH83} {very_unique} 3.5 Degradation of proteins, peptides and glycopeptides (26 Rco/19 Rpr) *rco:RC0603 acylamino-acid-releasing enzyme *rco:RC0228 thermostable carboxypeptidase *rco:RC0370 unknown (PARTIAL) similarity to ATP-dependent protease subunit C (ClpC) *rco:RC0727 unknown (PARTIAL) similarity to lysozyme {VERY_UNIQUE} *rco:RC1298 unknown (PARTIAL) similarity to lysozyme {VERY_UNIQUE} 4. Transporters 4.1 ABC transporters/amino acids (6 Rco/6 Rpr) 4.2 ABC transporters/nucleosides and nucleotides (3 Rco/3 Rpr) ABC transporters/cations (3 Rco/3 Rpr)) 4.3 4.4 ABC transporters/other (10 Rco/5 Rpr)) *rco:RC0414 unknown (PARTIAL) split gene of ABC transporter ATP-binding protein {SPLIT: RC0413 + RC0414 = bbu:BB0742} *rco:RC0413 unknown (PARTIAL) split gene of ABC transporter ATP-binding protein {SPLIT: RC0413 + RC0414 = bbu:BB0742} *rco:RC1187 uup; ABC transporter ATP-binding protein Uup (two ATP-binding cassettes) {DEGRADATION} 4.5 Protein and peptide secretion (14 Rco/14 Rpr) 46 Amino acids (19 Rco/10 Rpr) *rco:RC0832 unknown (PARTIAL) similarity to proline/betaine transporter *rco:RC1297 unknown (PARTIAL) similarity to proline/betaine transporter {DEGRADATION} *rco:RC0889 unknown (PARTIAL) similarity to proline/betaine transporter {SPLIT: RC0889 + RC0890 = RP313, RC0427 is the intact gene} *rco:RC0890 unknown (PARTIAL) similarity to proline/betaine transporter {SPLIT: RC0889 + RC0890 = RP313, RC0427 is the intact gene} *rco:RC1100 unknown (PARTIAL) split gene of proline/betaine transporter {SPLIT: RC1100 + RC1101 + RC1102 = RP313, RC0427

is the intact gene} *rco:RC1101 unknown (PARTIAL) split gene of proline/betaine transporter {SPLIT: RC1100 + RC1101 + RC1102 = RP313,

RC0427 is the intact gene}

Table 1 (continued)

4.7

*rco:RC1102 unknown (PARTIAL) split gene of proline/betaine transporter {SPLIT: RC1100 + RC1101 + RC1102 = RP313, RC0427 is the intact gene}
Nucleosides and nucleotides (9 Rco/5 Rpr)
*rco:RC0450 unknown (PARTIAL) similarity to DNA processing protein Smf {DEGRADATION}

*rco:RC0017 unknown (PARTIAL) similarity to DIVA processing protein simi {DEGRADATION}

*<u>rco:RC0843</u> unknown (PARTIAL) similarity to competence nocus E protein 5 {DEGRADATION} *<u>rco:RC0843</u> unknown (PARTIAL) split gene of competence protein F {SPLIT: RC0843 + RC0844 = hin:HI0434} {DEGRADATION}

*<u>rco:RC0844</u> unknown (PARTIAL) split gene of competence protein F {SPLIT: RC0843 + RC0844 = hin:H10434} {DEGRADATION}

4.8 Carbohydrates, organic alcohols and acids (1 Rco/1 Rpr)

4.9 Cations (7 Rco/2 Rpr)

*rco:RC0943 corA; magnesium and cobalt transport protein CorA {DEGRADATION}

*rco:RC0092 unknown (PARTIAL) similarity to cation efflux system protein {DEGRADATION}

*rco:RC0231 unknown (PARTIAL) similarity to cation transport regulator ChaB {DEGRADATION}

*rco:RC1354 unknown (PARTIAL) split gene of Na(+)/H(+) antiporter {SPLIT: RC1354 + RC1355 = cje:Cj1655c}

*<u>rco:RC1355</u> unknown (PARTIAL) split gene of Na(+)/H(+) antiporter {SPLIT: RC1354 + RC1355 = cje:Cj1655c} {DEGRADATION}

4.10 Other (5 Rco/5 Rpr)

5. Regulatory functions

5.1 Two-component system (9 Rco/7 Rpr)

5.2 Other regulatory functions (13 Rco/4 Rpr)

*rco:RC0063 similarity to transcriptional regulator {DEGRADATION}

*rco:RC0888 unknown (PARTIAL) similarity to (p)ppGpp 3-pyrophosphohydrolase

*rco:RC0374 unknown (PARTIAL) similarity to (p)ppGpp 3-pyrophosphohydrolase

*rco:RC0426 unknown (PARTIAL) similarity to (p)ppGpp 3-pyrophosphohydrolase {DEGRADATION}

*rco:RC1017 unknown (PARTIAL) similarity to ecotin precursor

*rco:RC1098 unknown (PARTIAL) split gene of (p)ppGpp 3-pyrophosphohydrolase {SPLIT: RC1098 + RC1099 = uur:UU283}

*rco:RC1099 unknown (PARTIAL) split gene of (p)ppGpp 3-pyrophosphohydrolase {SPLIT: RC1098 + RC1099 = uur:UU283}

*<u>rco:RC0907</u> unknown (PARTIAL) split gene of tryptophan repressor binding protein {SPLIT: RC0907 + RC0908 = xfa:XF1733} {VERY_UNIQUE}

*<u>rco:RC0908</u> unknown (PARTIAL) split gene of tryptophan repressor binding protein {SPLIT: RC0907 + RC0908 = xfa:XF1733} {VERY_UNIQUE}

6. Energy metabolism

6.1 Citrate cycle (15 Rco/15 Rpr)

6.2 Pyruvate metabolism (9 Rco/7 Rpr)

6.3 Oxidative phosphorylation (46 Rco/46 Rpr)

6.4 Fructose and mannose metabolism (4 Rco/3 Rpr)

*rco:RC1039 unknown (PARTIAL) split gene of mannose-1-phosphate guanylyltransferase {SPLIT: RC1039 + RC1040 = pae:PA3551}
*rco:RC1040 unknown (PARTIAL) split gene of mannose-1-phosphate guanylyltransferase {SPLIT: RC1039 + RC1040 = pae:PA3551}
*rpr:RP340 mannosyltransferase {VERY_UNIQUE}

7. Fatty acid and phospholipid metabolism

7.1 Fatty acid and phospholipid metabolism (38 Rco/26 Rpr)

*rco:RC0906 scoB; succinyl-CoA:3-ketoacid-coenzyme A transferase subunit B [EC:2.8.3.5] {DEGRADATION}

*rco:RC0556 similarity to glycerol-3-phosphate cytidyltransferase (tagD) {DEGRADATION}

*rco:RC0078 unknown (PARTIAL) similarity to BioC {DEGRADATION}

*<u>rco:RC0905</u> unknown (PARTIAL) split gene of succinyl-CoA:3-ketoacid-coenzyme A transferase {SPLIT: RC0904 + RC0905 = dra:DRA0068}

*<u>rco:RC0904</u> unknown (PARTIAL) split gene of succinyl-CoA:3-ketoacid-coenzyme A transferase {SPLIT: RC0904 + RC0905 = dra:DRA0068}

*<u>rpr:RP820</u> phbC2; poly-β-hydroxybutyrate polymerase {DEGRADATION}

Nucleotide metabolism Purine metabolism (9 Rco

Purine metabolism (9 Rco/7 Rpr)

*<u>rco:RC0126</u> unknown (PARTIAL) split gene of ribose-phosphate pyrophosphokinase {SPLIT: RC0125 + RC0126 = pae:PA4670} {DEGRADATION}

*<u>rco:RC0125</u> unknown (PARTIAL) split gene of ribose-phosphate pyrophosphokinase {SPLIT: RC0125 + RC0126 = pae:PA4670} {DEGRADATION}

8.2 Pyrimidine metabolism (8 Rco/8 Rpr)

8.3 Nucleotide sugars metabolism (1 Rco/1 Rpr)

9. Amino acid metabolism

- 9.1 Amino acid biosynthesis (3 Rco/2 Rpr) *rco:RC0139 argB; acetylglutamate kinase [EC:2.7.2.8] {DEGRADATION}
- 9.2 Amino acid degradation (5 Rco/4 Rpr) *rco:RC0621 gabD; succinate semialdehyde dehydrogenase [EC:1.2.1.16] {DEGRADATION}

10. Metabolism of cell envelop

10.1 Diaminopimelate (8 Rco/8 Rpr)

Table 1 (continued)

10.2	Peptideglycan biosynthesis (22 Rco/19 Rpr) * <u>rco:RC0249</u> pbpC; bifunctional penicillin-binding protein 1C * <u>rco:RC0318</u> unknown (PARTIAL) split gene of D-alanyl-D-alanine dipeptidase {SPLIT: RC0317 + RC0318 = sp:VANX_ENTFC} * <u>rco:RC0317</u> unknown (PARTIAL) split gene of D-alanyl-D-alanine dipeptidase {SPLIT: RC0317 + RC0318 = sp:VANX_ENTFC} Mombranes and linearestic (8 Res(8 Par)
10.3	Surface polysaccharides, lipopolysaccharides, glycerolipid and antigens (35 Rco/33 Rpr) * <u>rco:RC1273</u> rompA; 190-KDa cell surface antigen {DEGRADATION} * <u>rco:RC1261</u> similarity to cell surface antigen * <u>rco:RC1060</u> unknown (PARTIAL) similarity to lipopolysaccharide biosynthesis protein {VERY_UNIQUE} * <u>rco:RC1058</u> unknown (PARTIAL) split gene of α -(1,3)-fucosyltransferase {SPLIT: RC1058 + RC1059 = hpj:jhp1002} {DEGRADATION} * <u>rco:RC1059</u> unknown (PARTIAL) split gene of α -(1,3)-fucosyltransferase {SPLIT: RC1058 + RC1059 = hpj:jhp1002} {DEGRADATION}
11. 11.1	Metabolism of cofactors and other substances Folate biosynthesis (9 Rco/4 Rpr) *rco:RC0034 folA; dihydrofolate reductase [EC:1.5.1.3] {DEGRADATION} *rco:RC0037 whown (PARTIAL) split gene of folate synthesis bifunctional protein {SPLIT: RC0035 + RC0036 + RC0037 + RC0038 = ctr:CT613} *rco:RC0035 whown (PARTIAL) split gene of folate synthesis bifunctional protein {SPLIT: RC0035 + RC0036 + RC0037 + RC0038 = ctr:CT613} *rco:RC0036 unknown (PARTIAL) split gene of folate synthesis bifunctional protein {SPLIT: RC0035 + RC0036 + RC0037 + RC0038 = ctr:CT613} *rco:RC0036 unknown (PARTIAL) split gene of folate synthesis bifunctional protein {SPLIT: RC0035 + RC0036 + RC0037 + RC0038 = ctr:CT613} *rco:RC0038 unknown (PARTIAL) split gene of folate synthesis bifunctional protein {SPLIT: RC0035 + RC0036 + RC0037 + RC0038 = ctr:CT613}
11.2	Haem and porphyins metabolism (8 Rco/8 Rpr)
11.3	Ubiquinone/menaquinone biosynthesis (8 Rco/9 Rpr)
11.4	Lipoate (2 Rco/2 Rpr)
11.5	Inforedoxin and glutaredoxin (6 KC0/6 Kpr) Other (1 Reg/0 Rpr)
11.0	*rco:RC0033 unknown (PARTIAL) similarity to coenzyme PQQ synthesis protein c {DEGRADATION}
12.	Cellular processes
12.1	Cell division (16 Rco/15 Rpr)
12.2	*rco:RC0529 similarity to cell filamentation proteins (fic) {DEGRADATION}
12.2	rost-pathogen interaction (/ KCO/O Kpr)
12.3	Chaperones and stress-induced proteins (14 Rco/14 Rpr)
10	
13.	Drug and analogue sensitivity
13.1	*rco:RC1358 ampD: ampD protein homolog {DEGRADATION}
	*rco:RC0535 ampG; ampG protein {DEGRADATION}
	* <u>rco:RC0301</u> multidrug resistance ABC transporter ATP-binding protein {DEGRADATION}
	* <u>rco:RC0554</u> similarity to acetyltransferases
	*rco:RC0788 unknown (PARTIAL) similarity to MazG
	*rco:RC0947 unknown (PARTIAL) similarity to aminoglycoside 3'-phosphotransferase {VERY_UNIQUE}
	*rco:RC1181 unknown (PARTIAL) similarity to multidrug resistance protein {DEGRADATION}
	* <u>rco:RC1244</u> unknown (PARTIAL) split gene of β -lactamase {SPLI1: RC1243 + RC1244 = sp:BLAD_KLEPN} {DEGRADATION}
	* <u>ICO.RC1245</u> unknown (PARTIAL) split gene of probable RND efflux transporter {SPLIT:
	RC0564 + RC0565 + RC0566 + RC0567 + RC0568 + RC0569 = pae:PA2527
	*rco:RC0565 unknown (PARTIAL) split gene of probable RND efflux transporter {SPLIT:
	RC0564 + RC0565 + RC0566 + RC0567 + RC0568 + RC0569 = pae:PA2527}
	*rco:RC0567 unknown (PARTIAL) split gene of probable RND efflux transporter {SPLIT:
	RC0564 + RC0565 + RC0566 + RC0567 + RC0568 + RC0569 = pae:PA2527
	* <u>rco:KCU309</u> unknown (PAK11AL) split gene of probable KND emux transporter {SPL11: $P_{C0564} + P_{C0565} + P_{C0566} + P_{C0567} + P_{C0568} + P_{C0569} - n_{26} P_{A2527}$
	*rco:RC0568 unknown (PARTIAL) split gene of probable RND efflux transporter {SPLIT:
	$RC0564 + RC0565 + RC0566 + RC0567 + RC0568 + RC0569 = pae: PA2527 {DEGRADATION}$
	* <u>rco:RC0566</u> unknown (PARTIAL) split gene of probable RND efflux transporter {SPLIT:
	RC0564 + RC0565 + RC0566 + RC0567 + RC0568 + RC0569 = pae:PA2527} {DEGRADATION}
14.	Other categories
14.1	Adaptations to atypical conditions (11 Rco/11 Rpr)
14.2	Colcin-related functions (3 Rco/3 Rpr)

- 14.3 Uncategorized (46 Rco/37 Rpr)
 - *<u>rco:RC0860</u> ankyrin like protein *<u>rco:RC1090</u> bcp; bacterioferritin comigratory protein {DEGRADATION}

)	
*rco:RC0705 mccF2; similar to microcin C7 self-immunity protein {DEGRADATION}	
*rco:RC1093 putative periplasmic protein {DEGRADATION}	
*rco:RC0741 similarity to histon and other protein acetyltransferase	
*rco:RC0638 similarity to late-developmental spore coat protein	
*rco:RC0470 similarity to spore photoproduct lyase suhB	

*rco:RC0781 unknown (PARTIAL) similarity to N-terminal of biotin-protein ligase {VERY_UNIQUE}

*<u>rco:RC1364</u> unknown (PARTIAL) split gene of oxidoreductase {SPLIT: RC1363 + RC1364 = bsu:yxjF} {DEGRADATION}

*rco:RC1363 unknown (PARTIAL) split gene of oxidoreductase {SPLIT: RC1363 + RC1364 = bsu:yxjF} {DEGRADATION} *rpr:RP714 ank2; ankyrin, brain variant 2

Homology searches were carried out using the BLAST programs [22]. From both reconstructed and absent metabolic pathways of *Rickettsia*, as well as from specific characteristic of the strains compared, we tried to extrapolate a reconstruction of the different strategies invented by these bacteria to survive, replicate and make us ill.

3. Major metabolite pathway analysis

3.1. Carbohydrate metabolism

Metabolic enzymes are required to break down different carbon sources to yield energy. Bacteria can use different types of compounds as carbon sources, including sugars. Sugar metabolism starts with the glycolytic ("Embden-Meyerhof-Parnas") and Entner-Doudoroff pathways in many bacteria.

Genome sequence analyses revealed that enzymes for these pathways are totally absent in R. conorii and R. prowazekii (MAP00010, MAP00630). This is consistent with the failure to detect glycolytic activity from bacterial extracts [23]. Thus pyruvate, one of the major glycolysis products, cannot be produced by these pathways. However, the requirement of pyruvate for Rickettsia is apparent, given a large set of enzymes involved in metabolism of pyruvate (MAP00620) encoded in their genomes. For example, Rickettsia possess pyruvate dehydrogenase (pdhA, EC 1.2.4.1) and malate dehydrogenase (tme, EC 1.1.1.40). Pyruvate could also fuel the citrate cycle (TCA cycle or Krebs cycle) which was shown to be complete in *Rickettsia* (MAP00020). Pyruvate may be acquired directly from the host or in the form of phosphoenolpyruvate which may be converted to pyruvate by phosphate dikinase (EC 2.7.9.1). Acetoacetyl-CoA reductase (EC 1.1.1.36) is present in both R. conorii and R. prowazekii, probably for polyhydroxyalkanoic acid (PHA). PHA accumulates during bacterial survival in nutrient-depleted cultures [24]. Such an energy and carbon reservoir should favor Rickettsia survival upon nutrient starvation or entry into the stationary phase.

In contrast to *R. prowazekii*, the *R. conorii* genome encodes a succinate semi-aldehyde dehydrogenase homologue (MAP00650; *gabD*, EC 1.2.1.16). In other bacteria, *gabD* is not isolated in the chromosome but is part of a gene cluster encoding enzymes for γ -aminobutyrate (GABA) degradation [25]. Rickettsial genes for 4-aminobutyrate aminotransferase (*gabG* and *gabT*, EC 2.6.1.19) are not yet identified. The ability of rickettsiae to directly use GABA as a nitrogen source remains to be investigated.

The glyoxylate pathway (MAP00630) is absent in both *R. conorii* and *R. prowazekii*. This pathway is induced upon phagocytosis of some microorganisms in response to the nutrient starvation in the phagosome [26]. Rickettsiae rapidly escape from this compartment by dissolving the phagocytic vesicle surrounding them [27]. Thus, the utilization of this simple carbon source might not be necessary for these bacteria to survive in the host cell cytoplasm. Similarly, neither propanoate nor C5-branched dibasic acid can be used as sources of carbon and energy by *R. conorii* and *R. prowazekii* (MAP00640, MAP00660).

3.2. Energy metabolism

ATP is the universal biological energy currency. Aerobic microorganisms can synthesize ATP by two routes, namely glycolysis, which is lacking in Rickettsia as mentioned above, and oxidative phosphorylation [28]. The latter pathway (MAP00190) is composed of a respiratory chain with three proton pumps (NADH dehydro- genase = complex I, cytochrome c reductase or cytochrome bc_1 complex = complex III, and cytochrome c oxidase = complex IV), an enzyme complex not associated with proton translocation (succinate dehydrogenase = complex II), and the ATP synthase (complex V), a fourth reversible proton pump that takes advantage of the concentration gradient of protons across the membrane to synthesize ATP. The majority of the genes encoding the complexes are present in Rickettsia. Accordingly, the complete R. prowazekii genome sequence [14] provides evidence that mitochondria and

R. prowazekii exhibit very similar functional profiles with respect to ATP production (MAP00193) and reflects the common evolutionary origin of their electron transport chains. Both mitochondria and rickettsiae possess ATP/ADP translocases [29]. In contrast with mitochondria which export ATP to the cytosol, rickettsiae take up ATP produced by the hosts. The presence of a nucleotide transport system rather than a leaky membrane with exaggerated passive permeability was initially proposed in 1976 [30]. R. prowazekii and R. conorii exhibit 5 copies of the ATP/ADP translocase genes in their genome, further underscoring the importance of this system in their physiology. In fact, when rickettsiae enter the eukaryotic cell they first exploit the ATP already present in the host cell cytoplasm via ATP/ADP translocases. When the pool of the host ATP has been consumed, rickettsiae start to produce ATP via their own aerobic respiration.

Nitrogen metabolism is also lacking in both strains (MAP00910). When considering other obligate intracellular bacteria, we observed that glutamine synthetase (EC 6.3.1.2), an enzyme of central importance which catalyses the formation of glutamine from ammonia and glutamate [31] is present in Coxiella burnetti and Mycobacterium leprae but is lacking in Rickettsia as well as in Treponema pallidum. The absence of a recognizable gene for glutamine synthetase implies glutamine uptake from the host cells. This is likely to involve *glnP* encoding putative glutamine transport system permease protein in R. prowazekii and its orthologue in R. conorii, i.e., yqiY encoding an amino acid ABC transporter permease protein. When cultured within eukaryotic cells, the concentration of exogenously added glutamine in culture medium is higher for R. prowazekii (2%) than for R. conorii (1%) [32]. This difference is likely to result from a more efficient transport system for R. conorii.

3.3. Lipid metabolism

The majority of enzymes involved in fatty acids biosynthesis pathways (MAP00061 and MAP00062) are present. An important exception is acetyl-CoA carboxylase (EC 6.4.1.2), catalyzing the carboxylation of acetyl-CoA to malonyl-CoA, the first step of fatty acid synthesis. Acetyl-CoA carboxylase is an ubiquitous enzyme [33] in all genomes analyzed except for *Archeae*, in which isoprenoid lipids replace alkyl chain lipids. This enzyme is also lacking in *M. leprae* and *Tropheryma whipplei*. A common feature among these microorganisms is their very slow replication rate. *Rickettsia* may preferentially use the reverse reactions of the β -oxidation system (MAP00062) as in mitochondria.

3.4. Nucleotide metabolism

Nucleotide metabolism is divided into two pathways (purine and pyrimidine nucleotide metabolism, MAP00230, MAP00240). These nucleotides are both important as energy source and as RNA and DNA building blocks.

3.4.1. Purine metabolism

De novo synthesis of purines (adenine = 6-aminopurine; guanine = 2-amino-6-hydroxypurine; hypoxanthine = 6-oxy purine; xanthine = 2,6-dioxy purine) is not achieved by *Rickettsia*. The sole conserved enzyme is the 5-aminoimidazole-4-*N*-succinylcarboxamide ribonucleotide (SAICAR) synthetase encoded by *purC* (EC 6.3.2.6). In contrast, their genomes encode several enzymes for salvaging adenine and guanine. The lack of enzymes for the inter-conversion between adenine and guanine suggests that these bacteria depend on the host for both important purines.

The guanosine nucleotides pppGpp (guanosine 3'diphosphate 5'-triphosphate) and ppGpp (guanosine 3'-diphosphate 5'-diphosphate) are potent effectors of the bacterial stringent response [34]. Their accumulation, observed in response to nutritional starvation, initiates a global change in the cellular metabolism in E. coli. It is regulated in the cytosol by two enzymes, RelA (EC 2.7.6.5) and SpoT (EC 3.1.7.2). RelA is a synthase whereas SpoT is primarily a hydrolase that may exhibit a synthase activity under certain conditions [35]. Although full-length genes for these specific enzymes were not found in Rickettsia genomes, 5 short ORFs exhibiting significant amino acid sequence similarities (*E*-values $0.01 \sim 10^{-15}$, percent identities 29-60%) with SpoT/RelA homologues from other bacteria (e.g., Bradyrhizobium japonicum) were identified.

3.4.2. Pyrimidine metabolism

The de novo pyrimidine synthetic pathway, a cascade of reactions leading to orotic acid (2,4-dioxy-6-carboxy pyrimidine), is totally missing in *Rickettsia*. This contrasts with the conserved synthetic pathway of *T. whipplei* allowing it to grow under axenic conditions without pyrimidine supplementation [36]. This bacterium is the causative agent of Whipple's disease, a rare multisystem chronic infection involving the intestinal tract as well as various other organs [37]. From its whole-genome sequence analysis we identify specific metabolic deficiencies and design an axenic media allowing the growth of this host-dependent bacteria in cell-free conditions [36]. This paradigm illustrates how insights inferred from genomic studies to design physiological medium based on undestanding cellular metabolism and will be used as a reference throught this review. For pyrimidines and pyrimidine nucleotides, rickettsiae again depend on the host through salvage pathway enzymes, including recently found flavin-dependent thymidylate synthase ThyX [38]. Another major enzyme lacking in *Rickettsia* when compared to *T. whipplei* is purine nucleoside phosphorylase (EC 2.4.2.1). This enzyme catalyzes the formation of uracil from deoxyuridine. However, in *R. conorii*, an alternative route to generate uracil involves cytosine deaminase (EC 3.5.4.1) which is absent in *T. whipplei*. We identified *R. prowazekii* gene RP831 (previously unannotated. as the cytosine deaminase gene in this organism).

Experiments performed on *R. prowazekii* [39] suggested that these bacteria have a ribonucleotide reductase that would allow the synthesis of the deoxyribonucleotides required for DNA synthesis from ribonucleotides rather than relying on a transporter. Genome analysis confirmed the presence of genes (*nrdA* and *nrdB*) encoding ribonucleotide reductase subunits in *Rickettsia*.

3.4.3. Nucleotide sugars metabolism

Nucleotide sugars (MAP00520, MAP00530) serve as donors for glycosyltransferase reactions and also as initial compounds for cell wall biosynthesis. Several enzymes were identified in *Rickettsia* for the synthesis of UDP-*N*-acetyl-D-glucosamine, a key initial substrate for the synthesis of lipopolysaccharide (LPS) and peptidoglycan. In addition, rickettsiae exhibit UDPglucose 6-dehydrogenase (EC 1.1.1.22) and dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133), two enzymes involved in the synthesis of extracellular polysaccharides.

3.5. Amino acid metabolism

As detailed by Andersson et al. [12], amino acid metabolism is deficient in Rickettsia. One of the key compounds in amino acid metabolism is glutamate. In Rickettsia, glutamate metabolism is lacking (MAP00251) with the exception of aspartate aminotransferase A (aatA, EC 2.6.1.1) which is responsible for converting 2-oxoglutarate into L-glutamate. This amino acid is the precursor of both proline and ornithine. Proline biosynthesis starts with the ATP-driven phosphorylation and reduction of the carboxyl side chain of glutamate. It can also occur through the urea cycle. In this pathway, arginine serves as the first intermediate that is converted to ornithine. In *Rickettsia*, none of enzymes of the proline and arginine metabolism pathways (MAP00330) are present. Among all complete bacterial genomes, Chlamydia is the sole genus exhibiting exactly the same pattern. All others bacteria, including T. whipplei, Coxiella burnetti and Borrelia

burgdorferi are much better equipped with enzymes for amino acid metabolism.

Beside glutamate, aspartate is also important to collect and eliminate amino nitrogen via the urea cycle. In *Rickettsia*, aspartate and alanine metabolic pathways are largely absent (MAP00252). For several microorganisms, synthesis of D-alanine, which is a key component of peptidoglycan, can be catalyzed by alanine aminotransferase (EC 2.6.1.21) that produces D-alanine from pyruvate. This enzyme being absent in Rickettsia, D-alanine most likely originates from L-alanine isomerization (MAP00473) by alanine racemase (alr, EC 5.1.1.1). This amino acid can then be processed for peptidoglygan biosynthesis through D-alanine-D-alanine ligase (ddlB, EC 6.3.2.4). From aspartate we can also identify the initial steps for lysine biosynthesis while the diaminopimelate decarboxylase involved in the last step (EC 4.1.1.20) is absent (MAP00300). Thus the role of the conserved enzymes appears not for the lysine biosynthesis but rather for the synthesis of diaminopimwhich is an important component elate. of peptidoglycan [12]. The homoserine dehydrogenase (hom, thrA, EC 1.1.1.3), which catalyses L-homoserine synthesis and thus open the glycine, serine and threonine metabolism route is lacking (MAP00260). This corroborates experimental data showing that R. prowazekii growth depends on the host cell for serine or glycine [40]. In the same paper, the use of a host cell mutant provided in vivo evidence of a rickettsial enzymatic activity responsible for interconversion of both amino acids. A gene encoding for a serine hydroxymethyltransferase (glyA) was identified. Finally, rickettsiae are also deficient in histidine metabolism (MAP00340).Biosynthesis of the three branched-chain amino acids valine, leucine and isoleucine is absent (MAP00290). Their degradation can be initiated by the branched-chain amino acid aminotransferase (EC 2.6.1.42) encoded by *ilvE* which catalyzes a transamination reaction. This gene is not clustered with other genes of the branched-chain amino acid pathways (namely *leuB*, *leuCD*, *ilvB*, *ilvC* and *ilvD*) that are absent.

The first step in the biosynthesis of all three aromatic amino acids (tryptophan, tyrosine and phenylalanine) involves phosphoenolpyruvate (PEP) from glycolysis and erythrose-4-phosphate from the pentose phosphate pathway (MAP00400). These two precursors are transformed in shikimate as a metabolic intermediate, giving in turn chorismate. In *Rickettsia*, all the corresponding genes are missing and the synthesis of these amino acids thus seems impossible. Both phenylalanine and tryptophane metabolism are absent (MAP00360, MAP00380). As these aromatic amino acids are essentials for protein synthesis, they must be provided by the host cell or the culture medium. The same situation probably prevails for sulfur containing amino acids (methionine, cystine, cysteine, and taurine). In other bacteria, the cysteine sulfur comes from methionine metabolism, which is missing in Rickettsia (MAP00271). A condensation of ATP and methionine catalyzed by methionine adenosyltransferase (metK, EC 2.5.1.6) yields S-adenosylmethionine (AdoMet), that is then used as an essential methyl donor in various methyltransferase reactions. However, metK is present as a pseudogene status in different rickettsial genomes. The apparent inability of rickettsiae to synthesize Ado-Met is compensated by the presence of a rickettsial Ado-Met transporter, the first bacterial AdoMet transporter identified [41].

3.6. Glutathione metabolism

Glutathione metabolism is associated with oxygenic cyanobacteria and the oxygen-utilizing purple bacteria, but is absent in many other prokaryotes as is the case concerning Rickettsia (MAP00480). Novel thiols were described in bacteria lacking glutathione [42]. These compounds can be involved in prokaryotic sulfide metabolism, in the anaerobic production of CH_4 , or have an antioxidant role similar to glutathione. Information from bacterial genome sequences suggests that glutathione S-transferases are present in large numbers in proteobacteria [43]. Glutathione S-transferases constitute a large family of enzymes catalyzing the addition of glutathione to endogenous or xenobiotic, often toxic, electrophilic chemicals. Putative glutathione S-transferase (EC 2.5.1.18) genes were readily found in both R. prowazekii and R. conorii (RC0782, RP490).

3.7. Metabolism of complex carbohydrates

Starch and sucrose metabolism are absent (MAP00500). An isolated gene encoding a protein with similarity to β -glucosidase (EC 3.2.1.21) is present in the R. conorii genome, as well as its orthologue in R. prowazekii (RP706). These enzymes hydrolyze β-glycosidic linkages between the anomeric carbon and glycosidic oxygen in β -glucosides. They are peripheral membrane protein playing an important role in bacterial cellulose production [44]. In Rickettsia, the bglxA gene was found within an operon including rompB (outer membrane protein B or cell surface antigen sca5); htrB (EC 2.3.1.-) (lipid A biosynthesis lauroyl acyltransferase) and lpxK(EC 2.7.1.130) (tetraacyldisaccharide 4'-kinase). All these genes are involved in the metabolism of the bacterial envelop.

3.7.1. Lipopolysaccharide biosynthesis

Lipopolysaccharides (LPS) of Gram-negative bacteria are complex molecules also known as endotoxin, composed of Lipid A, core polysaccharide and O-polysaccharide [45]. Lipid A is a basal part of LPS, which is well conserved among different Gramnegative organisms. Covalently attached to the Lipid A is a core containing unusual and essential sugars of the Gram-negative bacterial LPS such as 3-deoxy-D-manno-octolusonic acid (KDO) O-polysaccharide. The O-chain is highly variable and confers serologic specificity. The presence of LPS within Rickettsia was first evidenced by SDS-PAGE analysis of whole cell extracts [46,47] and was shown to be group specific [48,49]. Its purification and characterization from the Thai tick typhus TT-118 and Katayama strains demonstrated the presence of KDO [50]. These experimental data are in accordance with the predicted biosynthetic LPS pathway (MAP00540) which is almost complete from UDP-N-acetylglucosamine to lauroyl-KDO2-lipid IV(A). Two missing enzymes in this pathway are the 3-deoxy- manno-octulosonate-8 phosphatase (EC 3.1.3.45) recently identified as encoded by *yrbI* gene in *E. coli* K12 [51] and the *lpxH*encoded protein (EC 3.6.1.-). The lpxH gene encodes the UDP-2,3-diacylglucosamine-specific pyrophosphatase that catalyzes the fourth step of lipid A biosynthesis in E. coli. About 50% of all Gram-negative organisms sequenced to date lack clear orthologs of lpxH. Hypothetical lpxH homologues with E values ranging from 0.003 to 0.2 (lpxH2) do exist in some bacteria lacking *lpxH*. Although most Gram-negative bacteria contain LpxH and/or LpxH2 homologues, at least seven organisms including R. conorii and R. prowazekii appear to lack both proteins. Others include Aquifex aeolicus, Chlamydophila pneumoniae, Chlamydia trachomatis, and Synechocystis sp [52]. It is suspected that bacteria lacking LpxH contain a distinct UDP-2,3diacylglucosamine-specific pyrophosphatase that remains to be identified. While LPS group specificity was experimentally demonstrated [48,49] the metabolic pathways of LPS synthesis are highly similar. One gene (RC1060) with homology to a LPS biosynthesis protein was exclusively identified in R. conorii.

3.7.2. Peptidoglycan biosynthesis

The peptidoglycan layer of bacterial cell walls contains a macromolecule composed of polysaccharide chains that are cross-linked by short peptide bridges. Peptidoglycan biosynthesis (MAP00550 and MAP00530) is an essential pathway for most bacteria [53]. Defects or disruption of the cell wall leads to cell lysis and death due to the high internal osmotic pressure of the cell. In rickettsial genomes we observed that the genes encoding enzymes involved in the cytoplasmic stage of peptidoglycan assembly (*murC*, *murD*, *murE*, *murF*, *ddlB*) are present. The only enzyme involved in the membrane stage is a translocase (phospho-*N*-acetylmuramoyl-pentapeptide-transferase, EC 2.7.8.13) encoded by *mraY*. The last step of petidoglycan

biosynthesis can be processed by the transferase encoded by murG (EC 2.4.1.-). In contrast to what was observed in *E. coli* [53] these genes are positioned in distinct clusters on the chromosome.

3.8. Metabolism of complex lipids

The sn-glycerol-3-phosphate acyltransferase (*plsB*, EC 2.3.1.15) which catalyzes the first step in membrane phospholipid formation is lacking (MAP00561). *E. coli* mutants defective in PlsB activity were isolated as glucose-3P (G3P) auxotrophs [54], this growth phenotype was attributed to the expression of an acyltransferase with an elevated K_m for G3P. Although rickettsial genomes exhibit predicted transporters (GlpT) for G3P, complementation of culture media with this compound may not be sufficient to synthesize glycerolipids when considering cell-free cultivation of these bacteria.

3.9. Metabolism of cofactors and vitamins

Many enzymatic activities require the presence of cofactors such as metal ions (Zn²⁺, Cu²⁺, Mn²⁺, K⁺, and Na⁺) or small organic molecules (coenzymes). Two of the most important coenzymes in the cell are nicotinamide adenine dinucleotide (NAD+) and its relative nicotinamide adenine dinucleotide phosphate (NADP+). While metabolism of their precursors (nicotinate and nicotinamine) are absent in Rickettsia, enzymes involved in the interconversion between NAD+ and NADP+ were identified (MAP00760). Thus, rickettsiae are endowed with a putative NAD+ kinase (EC 2.7.1.23) as well as the two subunits of the nicotinamide nucleotide transhydrogenase (EC 1.6.1.1, EC 1.6.1.2). This probably constitutes a part of mechanisms of interaction between rickettsiae and their host cells. Coenzyme A (CoA) is the major acyl group carrier in living systems. Its biosynthetic pathway, well defined in E. coli, consists of five enzymatic steps beginning with the vitamin pantothenate [55]. In Rickettsia, pantothenate can be transported into the cell by a sodium-dependent transporter. Two panF genes were found, one of them being split in R. conorii (RC0702 + RC0703 + RC0704 = RP465). Other enzymes are lacking with the exception of two enzymes around CoA (MAP00770).

A few bacteria do not make folates and must acquire them from external sources by using energy-requiring transport systems [56]. In *Rickettsia*, several steps are lacking for the biosynthetic pathway to tetrahydrofolate, the biologically active form, which is usually produced from GTP and chorismate (MAP00790). Thus, the step catalyzed by dihydropteroate synthetase (EC 2.5.1.15) is missing. This enzyme is the bacterial target of sulfamide antibiotic group and its lack account for the resistance of rickettsiae to cotrimoxazole (sulfamethoxazole + trimethoprim. [57]) Rickettsiae could have alternative enzymes for missing steps, or use transporters for folate or its derivatives.

In bacteria, biotin (vitamin H) serves as a co-factor for numerous biotin-dependent carboxylases and is synthesized by a well-defined pathway [58]. In *Rickettsia*, enzymes for the biosynthesis of this co-factor appear to be limited (MAP00780). Key genes are absent except *bioC* involved in an early step prior to pimeloyl CoA. The bi-functional BirA protein, both acting as a biotin-protein ligase and as a transcriptional repressor of the biotin operon in *E. coli* [59] is also present. In *R. conorii* this ubiquitous enzyme was found in the same operon as the transmembrane protein BioY involved in biotin transport and is likely to regulate biotin uptake.

About half of the ubiquinone biosynthetic pathway genes are present in *Rickettsia*, but almost all genes for menaquinone and enterochelin biosynthesis are missing (MAP00130). Genomes of *Rickettsia* contain homologs of *ubiA*, *ubiD*, *ubiX*, *ubiH*, *ubiE* and *ubiG*. In contrast, *ubiC*, *ubiB* and *ubiF* are missing. Cofactors often play essential roles in enzymatic reactions in bacteria. Some must be provided exogenously to certain bacteria as growth factors. From the above metabolic analyses of several cofactors, it is likely that *Rickettsia* acquires most of the cofactors or precursors from the host. It should be noted, however, that unidentified compensatory genes can achieve the functions of apparently absent enzymes and that some mistakes in sequencebased metabolic prediction are always possible.

4. Rickettsia membranes

One of the major constituents of the outer membrane of Gram-negative bacteria is lipopolysaccharide. As described previously, the biosynthetic pathway of this potent inducer of inflammation is complete in both R. conorii and R. prowazekii. The scrub typhus rickettsiae differ in that they lack LPS and belong in the separate although related genus, Orientia [60]. Other intracellular bacteria responsible for tick-borne diseases such as Ehrlichia chaffeensis and Anaplasma phagocytophilum are also devoid of LPS [61]. Few genes were identified for the biosynthesis of LPS within the Buchnera sp. strain APS [62], an endocellular symbiont harboured by pea aphids. This evolution, which results in a structural fragility, is likely to originate from the prolonged intracellular life of these bacteria, sheltered from attack by the host and foreign enemies.

Peptidoglycan cell wall biosynthesis is also complete in *Rickettsia*. The cell wall is the principal target of β -lactams antibiotics and their effectiveness is directly correlated with the ability to penetrate the cell either through passive diffusion or through porin channels. In contrast, glycoproteins (e.g., vancomycin) are unable to penetrate the outer membrane of Gram-negative bacteria and consequently have restricted activity against these organisms.

Rickettsia species differ not only in their LPS antigens but also in their surface-exposed proteins. A major immunodominant, 190-kDa surface exposed protein which has been named rOmpA (rickettsial Outer Membrane Protein A) is encoded by a gene (rompA) which exhibits several repeat units [63]. Genetic organization of these repeated regions was compared between different spotted fever group rickettsiae and their distinctive arrangements proposed as responsible for encoding protective species-specific conformational epitopes of the protein [64,65]. In addition, genetic variation of this gene allows the identification of various Rickettsia species based on PCR product sequencing [66]. SDS-PAGE and Western-blot analysis failed to demonstrated the presence of rOmpA in R. typhi and R. prowazekii [63] and the gene was never amplified by PCR. Comparative genomics highlighted the fact that this gene was initially present within R. prowazekii genome. As illustrated in Fig. 2, its evolution led to complete disappearance by degradation, leaving only traces of short homologous sequences in R. prowazekii. Involvement of this outer membrane protein in the adhesion of R. rickettsii to host cells was demonstrated a few years ago [67]. Based on the previous finding, a similar role in R. prowazekii and most problably in other typhus group rickettsiae has to be excluded. Another gene coding for a 70 kDa outer membrane protein (RC1261) was found exclusively in R. conorii. Other cell surface antigen genes (sca) are in various stages of evolution. This is the case for scal (RC0019) and sca4 (geneD, RC0667) which are split within the R. prowazekii genome. Indeed, these R. conorii ORFs are divided in two and three colinear ORFs in R. prowazekii, respectively (RC0019 = RP016 +RP017 + RP018 and RC0667 = RP498 + RP499). The gene sca2 (RC0110) is of 5,388 nucleotides in R. conorii while its homologue in R. prowazekii (RP081) is of only 1,026 nucleotides. The opposite scheme is also observed: sca3 is intact in R. prowazekii, but split in R. conorii $\{RC0630 + RC0631 = RP451\}$. Finally, the gene encoding rOmpB, the most abundant protein in the outer membrane of *Rickettsia*, is conserved (termed rompB in R. conorii and sca5 in R. prowazekii). It is also the case concerning other annotated outer membrane proteins encoded by omp1 (outer membrane protein Omp1), tolC (outer membrane protein TolC precursor), *nlpD* (outer membrane antigenic lipoprotein B precursor) and by the gene RC0105 homologous to RP075.

Relatively little is known regarding the rickettsial inner membrane structure. Genome analysis revealed the presence of a few set of genes susceptible to encode inner-membrane proteins. Among putative candidates is *yidC* recently involved in the biogenesis of respiratory chain complexes in E. coli [68]. SecE and SecG are two other inner membrane proteins involved in the Sec protein secretion pathway in addition to SecY. The later is believed to be the channel through which the translocating polypeptide passes across the inner membrane [69]. Another putative rickettsial integral membrane protein is encoded by the gene RC1372 (ortholog RP883). It must be noted that a large majority of genes encode for proteins of unknown function, and most probably several specific membrane proteins belong in this category.



Fig. 2. Schematic representation of the rOmpA gene disappearance within R. prowazekii genome.

5.1. ABC transporters

Transporters of the ABC (ATP-binding cassette) class are encoded in large sets of paralogous genes in all bacterial species. An operon structure is frequently encountered for the three genes encoding each transporter: the ATP-binding protein, the membrane protein, and the substrate-binding protein [70]. This strategy used by Gram-negative bacteria allows not only the uptake of nutrients but also the secretion of small molecules and proteins lacking an N-terminal signal. Among these ABC protein-mediated exporters, we identified three distinct Ccm (Cytochrome C maturation) proteins (CcmA, CcmB and CcmC) previously classified as haem exporters [71]. Thus far several lines of evidence suggest that CcmAB is not a haem transporter [72] and its function remains to be elucidated. In bacteria, the ABC superfamily genes are mostly devoted to nutrient import. Of 22 genes present in Rickettsia, we found homologues of glnP and glnQ, two genes involved in glutamine transport in E. coli, reinforcing the importance of a glutamine-rich environment for Rickettsia growth.

5.2. Multidrug transporters

One of the mechanisms that bacteria use to circumvent the toxic effects of antibiotics is the active extrusion of these molecules from the cell through multidrug transporters [73]. Five super-families of bacterial multidrug efflux transporters primarily designed on the basis of amino acid sequence homology were described. These include the ATP-binding cassette (ABC) family composed by active transporters dependent on ATP hydrolysis and represented within *Rickettsia* genomes (abcT3, *msbA1*, *msbA2*). Other active transporters are driven by electrochemical ion gradients that exist across the plasma membrane. Rickettsia possess genes encoding for these proton-dependent multidrug efflux proteins and are classified into the major facilitator superfamily (MFS; bcr2), the small multidrug resistance family (SMR; emrA and emrB), the resistance/nodulation/cell division family (RND; aprE) and the multidrug and toxic compound extrusion family (MATE; mviN).

5.3. Protein and peptide secretion

Proteins that have to be transported across the bacterial plasma membrane are, in most cases, initially synthesized as larger precursor proteins containing an N-terminal signal peptide [74]. Three major categories of translocation across the cytoplasmic membrane have been described in bacterial cells: Sec-mediated translocation, Tat-mediated translocation, and other specialized mechanisms [75]. The Sec pathway is generally involved in the transport of newly synthesized proteins out of the cytosol before they acquire their final structures. This Sec machinery consists of a protein conducting channel (SecYEG) and the SecA translocation ATPase, which couples the energy of ATP binding and hydrolysis to protein translocation [76]. The complete set of corresponding genes was found in *Rickettsia* (*secA*, *secB*, *secD*, *secE*, *secF*, *secG*, *secY*).

The twin-arginine translocation (Tat) pathway represents an alternative possibility for posttranslational translocation of proteins across the cytoplasmic membrane [77]. The Tat system is different from the Sec pathway because of its unusual ability to transport folded proteins and even enzyme complexes into the periplasm. For *Rickettsia*, only one gene (*tatC*) encoding components of the Tat export machinery has been identified.

In eukaryotes, proteins synthesized in the cytosol can also be targeted to the endoplasmic reticulum via interactions with the signal recognition particle (SRP) which is a protein–RNA complex [78]. In *Rickettsia*, SRP is comprised solely of Ffh, the small RNA (4.5S RNA) being absent. This protein homologous to the eukaryotic SRP54 is supposed to bind to the N-terminal signal sequences of secreted proteins and to target them to the membrane in conjunction with FtsY.

The type IV secretion system machinery is a multicomponent pore that allows the delivery of virulence factors from bacteria across the bacterial and host membranes into the cytoplasm of the host cell [79]. Seven of the 12 genes annotated in E. coli are present in Rickettsia. Despite a significant difference in virB gene numbers and genetic loci, the type IV secretion machinery of both R. conorii and R. prowazekii is phylogenetically close to that of A. phagocytophilum and E. chaffeensis [80]. Identification of the nature of the molecules secreted by the type IV secretion system may help in understanding pathogenesis. Therefore, and in regards with rickettsial diseases, one of the most interesting outer membrane components of the type IV secretion system is VirB9. A recent study performed on Ehrlichia canis has evidenced the highly antigenic properties of this protein which is expressed both in mammalian and tick hosts [81]. This suggests that inhibition of VirB9 function may inhibit the intracellular survival of E. canis and thus may constitute a good vaccine candidate for canine ehrlichiosis. Finally, the rickettsial protein secretion system includes signal peptidases among which is the gene lepB which encodes a type I signal peptidase. These bacterial signal peptidases are unique serine proteases that utilize a Ser/Lys catalytic dyad mechanism in place of the classical Ser/His/Asp catalytic triad mechanism. They cleave the amino-terminal signal peptide extension from proteins that are translocated across biological membranes and represent a potential novel antibiotic target at the bacterial membrane surface [82]. The cloned *lepB* genes from *R. rickettsii* and *R. typhi* have been demonstrated to possess signal peptidase I activity [83].

5.4. Amino acid transporters

The ProP protein is an osmoregulatory transporter mediating the active accumulation of diverse compatible solutes, including proline and glycine betaine (*N*-trimethyl glycine). Uptake of osmoprotectants by ProP, activated by an osmotic upshift in *E. coli*, allows bacteria to survive drastic osmotic shifts [84]. In *Rickettsia*, this gene is present in multiple copies and is likely to be one of the adaptive mechanisms allowing these bacteria to survive and grow in diverse environments.

5.5. Nucleoside and nucleotide transporters

Like *Chlamydia*, *Rickettsia* are able to exchange their intracellular ADP for the host cell ATP [85]. Detailed molecular analysis of the R. prowazekii ATP/ADP translocase Tlc1 was described by Dunbar et al. [86]. The topological model of R. prowazekii ATP/ADP translocase is in agreement with the proposed in silico topological model in which Tlc traverses the cytoplasmic membrane of E. coli 12 times with its N- and C-terminal facing the cytoplasm [87]. R. conorii and R. prowazekii genome sequences each exhibit five copies of ATP/ ADP translocase homologues. Here again, such a high number might be the consequence of an adaptative response to unusual ecological niches. In general, Rickettsia evolved to possess transport systems allowing them to take full advantage of the many preformed metabolites abundant in the host cell, as is the case for ATP and the expansion of the Tlc repertoire.

6. Virulence factors

The search for the determinants of rickettsial virulence was initiated long ago through in vitro studies of *Rickettsia*-host cell interactions. As for other pathogens, these interactions involve several steps including recognition, entry, phagosome escape, growth, and cell lysis. The internalization of rickettsiae by their host cells requires adherence to an unidentified plasma membrane receptor by viable, metabolically active organisms [88]. Li and Walker [67] demonstrated that both anti-rOmpA monoclonal antibodies and the purified protein have a competitive inhibitory effect on the attachment of R. rickettsii to host cells, suggesting its role in the initial adhesion of the spotted fever group rickettsiae. Indeed, and as previoulsly discussed, this protein disappeared in the course of evolution and is no longer present in typhus group rickettsiae. A recent study suggested that the

rOmpB which is highly conserved, plays an important role in this recognition and consequently in invasion of host cells by rickettsiae [89]. Both outer membrane proteins are likely to play a key role in the interaction of these microorganisms with their host cells. Immunization of mice with fragments of both genes partially protect these animals against subsequent virulent challenge with spotted fever group rickettsiae [90]. Protection against a lethal R. conorii injection to mice was also afforded with antibodies against rOmpA and rOmpB [91]. To date, in vivo experiments have not been described with R. prowazekii or R. typhi. Uptake of Rickettsia also requires participation of the parasitized host as it is mediated by an induced phagocytosis process [92,93]. A rickettsial phospholipase A₂ (PLA₂) has been proposed to mediate escape from phagosomes [27,94]. PLA₂ activity was also associated in part with the hemolytic activity of *R. prowazekii* [95]. This well characterized property of R. prowazekii involves cholesterol as an essential component of the sheep erythrocyte receptor to which bacteria adsorb before lysing the cell [96]. It was therefore demonstrated that pure PLA₂ failed to have a hemolytic activity and that rickettsiae may elaborate cofactors beside phospholipase [97]. Once within the cytoplasm, rickettsiae are able to proliferate inside eukaryotic host cells, are free to exploit the nutrient-rich environment and to interact with host structural components. One of the characteristics of spotted fever group rickettsiae is their ability to exploit host actin pools [93,98,99]. In contrast to typhus group rickettsiae, spotted fever group rickettsiae move to occupy the entire cytoplasm of the infected cell and to spread between cells. The strict intracellular environment of Rickettsia introduces technical difficulties which have long prohibited their detailed study. The need to grow and extract these bacteria from cultured eukaryotic cells leads to possible inactivation, alteration, and contamination of these microorganisms. In addition, efforts to understand molecular aspects of the infection have been hampered by systematic difficulties, including the lack of genetic tools to manipulate Rickettsia [6]. Of note is that successful isolation of the fist knock-out of R. prowazekii was obtained only a few months ago [7]. This technical advance associated with completion of the sequencing of the genomes of both R. conorii and R. prowazekii should provide the opportunity to identify putative virulence factors within these strict intracellular pathogens. The exact definition of bacterial virulence factors has been widely discussed and various definitions of microbial pathogenicity have been proposed [100]. In essence, a virulence factor is any moiety produced by a pathogen that is essential for causing disease in a host. Proteins thought to be necessary for pathogenicity fall into several categories, according to their mechanism of virulence and function. When linked to lateral gene transfer events, virulence factors are found clustered within genomic regions called pathogenicity islands. Such genomic segments are easily recognizable through distinctive sequence features [101] but were not identified within any Rickettsia for which whole-genome sequencing has been completed. In fact, similarity searches using Rickettsia genomic sequences have pointed out only a few candidates including three hemolysin homologues. These putative virulence factors have been studied by complementation experiments. For instance, transferring R. typhi tlyC gene to the hemolysin-negative Proteus mirabilis hpmA mutant conferred it a hemolytic phenotype. From this result, a role for hemolysin in host cell entry, exit from the phagosome, and host cell lysis was hypothesized [102]. The transport of protein toxins from bacterial cytoplasm into the host or in the extracellular matrix is believed to be regulated by a type IV secretion system [79]. Among the substances that can be secreted through such a transporter are also two invasins, including the invA gene product [103]. Such proteins differ from adherence factors by acting extra-cellularly, breaking down host defenses at the local level and easing the passage of the infection. Identification of the phospholipase suspected to be involved in the pathogenesis of Rickettsia required an extensive analysis of genomic sequences. By using a sequence analysis program to identify all phospholipase signatures throughout the rickettsial genome, we identified an unique putative phospholipase protein belonging to the phospholipase D (PLD) superfamily. This enzyme might account for the activity previously attributed to PLA2 and could be critical to the intracellular life of these bacteria [104]. Finally, concerning R. conorii, the underlying mechanism of actin-based motility involves the protein RickA which acts through Arp2/3 recruitment and activation [105]. This protein is encoded by RC0909 now termed rickA. No ortholog was found within R. prowazekii which correlates with their lack of motility. Other rickettsial genes are most probably involved in host cell entry but remain to be identified.

7. Environmental adaptation of Rickettsia

7.1. DNA repair mechanisms

Because chemical stability of DNA is one of the prerequisites of life, DNA repair can be considered as a major defense system against environmental damages. Repair processes are divided in three major classes based on their general mechanisms of action [106,107] including direct repair, excision repair and recombinational repair. *Rickettsia* do not encode any protein involved in direct repair (abnormalities that are chemically reversed). Due to their intracellular nature, these bacteria are naturally protected from both UV- and alkylation-induced DNA lesions. Phylogenetic analysis of proteins involved in this process, namely photolyases and alkyltranferases, suggests that these ancient proteins were lost in course of evolution [107]. Excision repair includes several categories including base excision repair (BER), mismatch repair (MMR) and nucleotide excision repair (NER). Enzymes involved in BER are DNA glycosylases and endonucleases. DNA-3-methyladenine glycosidase (mpg, EC 3.2.2.21) is present in both strains. In contrast, formamidopyrimidine-DNA glycosidase (mutM, EC 3.2.2.23) and 7,8-dihydro-8-oxoguaninetriphosphatase (mutT, EC 3.6.1.-) which primarily act in the protection against mutations due to oxidative DNA damage [108] are exclusively found in R. conorii. Members of the Xth endonuclease family are also represented in Rickettsia. Essential MMR proteins (mismatch recognition proteins, exonucleases and helicase) were also identified (*mutS*, *mutL*, *recJ*, *xseA*, *xseB* and *uvrD*). The gene encoding the Dam protein is lacking as is the other well-described bacterial methyltranferase, CcmR. Finally, the NER system is complete in *Rickettsia* (uvrA, uvrB, uvrC, uvrD, mfd). This system which is associated with bacterial membranes is thought to have a transport function [107]. Recombinational repair is a system that uses homologous recombination to repair abnormalities. This repair pathway, in which the RecA protein (recombinase) plays a central role, is rather complex. The R. prowazekii recA gene coding a product able to complement recombinational deficiencies in E. coli mutants was experimentally demonstrated 10 years ago [109]. Genome sequencing allowed identification of additional genes involved in homologous recombination. These genes are recF, recJ, recN and recR (Rec-BCD initiation pathway), ruvA, ruvB, recG (branch migration), ruvC, recG (resolution step). Thus, Rickettsia seem to be well-equipped to resist to DNA damage.

7.2. Genomic evolution

Adaptation to changes in environmental conditions is indispensable. This is particularly true when considering the habitats of Rickettsia, which encompass a wide range of conditions. The main characteristic of these bacteria is their obligate intracellular nature. These microorganisms cannot survive outside the protected environment supplied by their host. This implicates the evolution of bacterial machinery geared toward the exploitation of the host biosynthetic pathways. Some of the molecular mechanisms likely to be central in host-pathogen interactions have been identified (amino acids transporters, ADP/ATP transporters, etc.). In the course of evolution, these bacteria also acquired genomic features to adapt to extreme environmental conditions such as high or cold temperatures within the ticks (R. conorii), pH changes in the surroundings, nutrient starvation. We believe that Rickettsia-specific

multi-copy genes also play a crucial role in these processes. Positive selection of such genes is likely, as the retention of multiple copies of genes is contrary to the general tendency of intracellular – small genome – bacteria that minimize genetic redundancy.

The gene *proP* is among the genes found in multiple copies. This transporter allows transfer of small molecules known as osmolytes including some amino acids, notably proline, which is osmoprotective for bacteria facilitating their growth in high salinity environments. Accumulation of proline in the cytoplasm is indeed accompanied by a reduction in the concentrations of less compatible solutes and an increase in cytosolic water volume [110]. More *proP* copies are present in *R. prowazekii* suggesting that such a mechanism promoting cellular rehydratation could favor survival of *R. prowazekii* within lice feces where they can survive for up to 100 days (unpublished observations). ADP/ATP transporters constitute another system likely to intervene in *Rickettsia* adaptation.

Adaptation to various conditions can also occur through another strategy, i.e., the coordinated expression/repression of specific genes. Such gene regulation is known to occur in response to environmental stress or during entry into stationary phase. When bacterial cells are deprived of an amino acid or carbon source, changes in many processes are seen within the cells. In E. coli and many other organisms, this phenomenon, called stringent response, is regulated by two unusual phosphorylated nucleotides derivatives of GTP and GDP, guanosine tetra- (ppGpp) and pentaphosphates (pppGpp), respectively [111]. In growth-favourable conditions, a steady-state level of these alarmones is maintained in the cytosol by two enzymes, RelA and SpoT. RelA is a synthase (also known as 3'-kinase), whereas SpoT is primarily a hydrolase that also has synthase activity under certain conditions. Interestingly Rickettsia genomes are endowed with 5 gene copies that have homology to spoT. The role of these genes in Rickettsia remains to be investigated. Preliminary data suggest that the transcription of at least one of the spoT homologues is correlated with nutritional stress (unpublished results).

8. Concluding remarks

The sequencing and annotation of *Rickettsia* have provided significant information on the basic nature of these bacteria including their core metabolism, secretion, adhesion, transporters, and cell-wall features. These intracellular bacteria have a degrading genome, following the paradigm observed with many host-associated organisms. Analysis of these genomes also highlights a genomic inventory in which some specific features such as exotic repeated elements [112] and a large number of "split" genes [13] are shared. Therefore, a genome sequence is not an end in itself. A major challenge still has to be met to improve understanding of how the genome functions. Once a gene has been identified the question of the activity of the potentially encoded protein has to be addressed. Experimental data are needed to complement and extend the results of homology studies. A conventional method is to assign a gene function by inactivation. Recent progress in the field of the genetic transformation of Rickettsia [7] would undoubtedly open the way to a better understanding of the molecular mechanisms involved in the pathogenicity of these bacteria. Information about the activity of a protein coded by a gene can also be provided by protein engineering. Here again, major difficulties are encountered when considering rickettsial proteins [113]. Another point of importance is that the genome sequences reveal large gaps in our knowledge. Of the 1374 ORFs in the R. conorii genome, 49% (37.3% in R. prowazekii). correspond to putative proteins of unknown function or without homologs in databases [12,13]. In this respect, we believe that another way to assign function to an unknown gene is by application of comparative genomics. As rewieved by Koonin [114], comparative genomics has an important role to play in genome function analysis because the presence or alternatively the absence of a homolog in another related organism can be a key to understand the role of a gene. Rickettsiae are potential bioterrorism agents for which vaccines are still not available. Analysis of R. conorii and R. prowazekii genomes has revealed only a few of the potential virulence factors usually associated with pathogens. However, as detailed in this review, comparative analysis of such a pair of syntenic genomes allowed identification of specific membrane proteins susceptible to be targeted for serological diagnosis. In short-term, availability of others rickettsial genomes for which sequencing is currently in progress, should permit the design of individual sets of strainspecific genes. The available genome sequences will now be used to guide post-genomics approaches (i.e. transcriptome analysis, proteomics, protein expression and functional studies) aiming at a better understanding of rickettsial virulence and physiology with the ultimate goal of improving diagnostics and therapeutics.

Acknowledgements

Authors acknowledge Pr. J.S. Dumler (The Johns Hopkins University School of Medicine, Baltimore, USA) for critical reading of the manuscript.

References

- Roux, V. and Raoult, D. (1995) Phylogenetic analysis of the genus *Rickettsia* by 16S rDNA sequencing. Res. Microbiol. 146, 385–396.
- [2] Raoult, D. and Roux, V. (1997) Rickettsioses as paradigms of new or emerging infectious diseases. Clin. Microbiol. Rev. 10, 694–719.
- [3] Rachek, L.I., Hines, A., Tucker, A.M., Winkler, H.H. and Wood, D.O. (2000) Transformation of *Rickettsia prowazekii* to erythromycin resistance encoded by the *Escherichia coli* ereB Gene. J. Bacteriol. 182, 3289–3291.
- [4] Troyer, J.M., Radulovic, S. and Azad, A.F. (1999) Green fluorescent protein as a marker in *Rickettsia typhi* transformation. Infect. Immun. 67, 3308–3311.
- [5] Renesto, P., Gouin, E. and Raoult, D. (2002) Expression of green fluorescent protein in *Rickettsia conorii*. Microb. Pathog. 33, 17–21.
- [6] Wood, D.O. and Azad, A.F. (2000) Genetic manipulation of *Rickettsiae*: a preview. Infect. Immun. 68, 6091–6093.
- [7] Qin, A., Tucker, A.M., Hines, A. and Wood, D.O. (2004) Transposon mutagenesis of the obligate intracellular pathogen *Rickettsia prowazekii*. Appl. Environ. Microbiol. 70, 2816–2822.
- [8] Tarasevich, I., Rydkina, E. and Raoult, D. (1998) Outbreak of epidemic typhus in Russia. Lancet 352, 1151.
- [9] Raoult, D., Ndihokubwayo, J.B., Tissot-Dupont, H., Roux, V., Faugere, B., Abegbinni, R. and Birtles, R.J. (1998) Outbreak of epidemic typhus associated with trench fever in Burundi. Lancet 352, 353–358.
- [10] Stein, A., Purgus, R., Olmer, M. and Raoult, D. (1999) Brill-Zinsser disease in France. Lancet 353, 1936.
- [11] Azad, A.F. and Radulovic, S. (2003) Pathogenic *Rickettsiae* as bioterrorism agents. Ann. N.Y. Acad. Sci. 990, 734–738.
- [12] Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H. and Kurland, C.G. (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133–140.
- [13] Ogata, H., Audic, S., Renesto-Audiffren, P., Fournier, P.E., Barbe, V., Samson, D., Roux, V., Cossart, P., Weissenbach, J., Claverie, J.M. and Raoult, D. (2001) Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. Science 293, 2093–2098.
- [14] Andersson, J.O. and Andersson, S.G.E. (1999) Genome degradation is an ungoing process in *Rickettsia*. Mol. Biol. Evol. 16, 1178–1191.
- [15] Ochman, H., Lawrence, J.G. and Groisman, E. (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405, 299–304.
- [16] Garcia-Vallvé, S., Romeu, A. and Palau, J. (2000) Horizontal gene transfer in bacterial and archaeal complete genomes. Genome Res. 10, 1719–1725.
- [17] Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M. (2000) Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39. Nucleic Acids Res. 28, 1397– 1406.
- [18] Koonin, E.V., Makarova, K.S. and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. Annu. Rev. Microbiol. 55, 709–742.
- [19] Mahan, M.J. and Roth, J.R. (1989) Role of recBC function in formation of chromosomal rearrangements: a two-step model for recombination. Genetics 121, 433–443.

- [20] Kanehisa, M., Goto, S., Kawashima, S. and Nakaya, A. (2002) The KEGG databases at GenomeNet. Nucleic Acids Res. 30, 42–46.
- [21] Paulsen, I.T., Nguyen, L., Sliwinski, M.K., Rabus, R. and Saier Jr., M.H. (2000) Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. J. Mol. Biol. 301, 75–100.
- [22] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
- [23] Coolbaugh, J.C., Progar, J.J. and Weiss, E. (1976) Enzymatic activities of cell-free extracts of *Rickettsia typhi*. Infect. Immun. 14, 298–305.
- [24] Matin, A., Veldhuis, C., Stegeman, V. and Veenhuis, M. (1979) Selective advantage of a *Spirillum sp.* in a carbon-limited environment. Accumulation of poly-β-hydroxybutyric acid and its role in starvation. J. Gen. Microbiol. 112, 349–355.
- [25] Schneider, B.L., Ruback, S., Kiupakis, A.K., Kasbarian, H., Pybus, C. and Reitzer, L. (2002) The *Escherichia coli* gabDTPC operon: specific gamma-aminobutyrate catabolism and nonspecific induction. J. Bacteriol. 184, 6976–6986.
- [26] Oh, M.K., Rohlin, L., Kao, K.C. and Liao, J.C. (2002) Global expression profiling of acetate-grown *Escherichia coli*. J. Biol. Chem. 277, 13175–13183.
- [27] Winkler, H.H. and Miller, E.T. (1982) Phospholipase A and the interaction of *Rickettsia prowazekii* and mouse fibroblasts (L-929 cells). Infect. Immun. 38, 109–113.
- [28] Schagger, H. (2002) Respiratory chain supercomplexes of mitochondria and bacteria. Biochim. Biophys. Acta 1555, 154– 159.
- [29] Winkler, H.H. and Neuhaus, H.E. (1999) Non-mitochondrial ATP transport. Trends Biochem. Sci. 24, 64–68.
- [30] Winkler, H.H. (1976) Rickettsial permeability. An ADP-ATP transport system. J. Biol. Chem. 251, 389–396.
- [31] Eisenberg, D., Gill, H.S., Pfluegl, G.M. and Rotstein, S.H. (2000) Structure–function relationships of glutamine synthetases. Biochim. Biophys. Acta 1477, 122–145.
- [32] Roux, V., Rydkina, E., Eremeeva, M. and Raoult, D. (1997) Citrate synthase gene comparison, a new tool for phylogenetic analysis, and its application for the *rickettsiae*. J. Syst. Bacteriol. 47, 252–261.
- [33] Cronan, J. and Waldrop, G.L. (2002) Multi-subunit acetyl-CoA carboxylases. Prog. Lipid Res. 41, 407–435.
- [34] Cashel, M. and Gallant, J. (1969) Two compounds implicated in the function of the RC gene of *Escherichia coli*. Nature 221, 838– 841.
- [35] Avarbock, D., Salem, J., Li, L.S., Wang, Z.M. and Rubin, H. (1999) Cloning and characterization of a bifunctional RelA/ SpoT homologue from *Mycobacterium tuberculosis*. Gene 233, 261–269.
- [36] Renesto, P., Crapoulet, N., Ogata, H., La Scola, B., Vestris, G., Claverie, J.M. and Raoult, D. (2003) Genome-based design of a cell-free culture medium for *Tropheryma whipplei*. Lancet 362, 447–449.
- [37] La Scola, B., Fenollar, F., Fournier, P.E., Altwegg, M., Mallet, M.N. and Raoult, D. (2001) Description of *Tropheryma whipplei* gen. nov., sp. nov., the Whipple's disease bacillus. Int. J. Syst. Evol. Microbiol. 51, 1471–1479.
- [38] Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P. and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. Science 297, 105–107.
- [39] Cai, J., Speed, R.R. and Winkler, H.H. (1991) Reduction of ribonucleotides by the obligate intracytoplasmic bacterium *Rickettsia prowazekii*. J. Bacteriol. 173, 1471–1477.

- [40] Austin, F.E., Turco, J. and Winkler, H.H. (1987) *Rickettsia prowazekii* requires host cell serine and glycine for growth. Infect. Immun. 55, 240–244.
- [41] Tucker, A.M., Winkler, H.H., Driskell, L.O. and Wood, D.O. (2003) S-Adenosylmethionine transport in *Rickettsia prowazekii*. J. Bacteriol. 185, 3031–3035.
- [42] Fahey, R.C. (2001) Novel thiols of prokaryotes. Annu. Rev. Microbiol. 55, 333–356.
- [43] Vuilleumier, S. and Pagni, M. (2002) The elusive roles of bacterial glutathione S-transferases: new lessons from genomes. Appl. Microbiol. Biotechnol. 58, 138–146.
- [44] Tajima, K., Nakajima, K., Yamashita, H., Shiba, T., Munekata, M. and Takai, M. (2001) Cloning and sequencing of the βglucosidase gene from *Acetobacter xylinum* ATCC 23769. DNA Res. 8, 263–269.
- [45] Erridge, C., Bennett-Guerrero, E. and Poxton, I.R. (2002) Structure and function of lipopolysaccharides. Microbes Infect. 4, 837–851.
- [46] Anacker, R.L., Philip, R.N., Williams, J.C., List, R.H. and Mann, R.E. (1984) Biochemical and immunochemical analysis of *Rickettsia rickettsii* strains of various degrees of virulence. Infect. Immun. 44, 559–564.
- [47] Teysseire, N. and Raoult, D. (1992) Comparison of Western immunoblotting and microimmunofluorescence for diagnosis of Mediterranean spotted fever. Infect. Immun. 30, 455–460.
- [48] Hechemy, K.E., Raoult, D., Eisemann, C., Han, Y.S. and Fox, J.A. (1986) Detection of antibodies to *Rickettsia conorii* with a latex agglutination test in patients with Mediterranean spotted fever. J. Infect. Dis. 153, 132–135.
- [49] Vishwanath, S. (1991) Antigenic relationships among the rickettsiae of the spotted fever and typhus groups. FEMS Microbiol. Lett. 65, 341–344.
- [50] Amano, K., Fujita, M. and Suto, T. (1993) Chemical properties of lipopolysaccharides from spotted fever group rickettsiae and their common antigenicity with lipopolysaccharides from Proteus species. Infect. Immun. 61, 4350–4355.
- [51] Wu, J. and Woodard, R.W. (2003) *Escherichia coli* YrbI is 3deoxy-D-manno-octulosonate 8-phosphate phosphatase. J. Biol. Chem. 278, 18117–18123.
- [52] Babinski, K.J., Kanjilal, S.J. and Raetz, C.R.H. (2002) Accumulation of the lipid A precursor UDP-2,3-diacylglucosamine in an *Escherichia coli* mutant lacking the lpxH Gene. J. Biol. Chem. 277, 25947–25956.
- [53] Azzolina, B.A., Yuan, X., Anderson, M.S. and El Sherbeini, M. (2001) The cell wall and cell division gene cluster in the Mra operon of pseudomonas aeruginosa: cloning, production, and purification of active enzymes. Protein Expr. Purif. 21, 393–400.
- [54] Bell, R.M. (1975) Mutants of *Escherichia coli* defective in membrane phospholipid synthesis. Properties of wild type and Km defective sn-glycerol-3-phosphate acyltransferase activities. J. Biol. Chem. 250, 7147–7152.
- [55] Rock, C.O., Park, H.W. and Jackowski, S. (2003) Role of feedback regulation of pantothenate kinase (CoaA) in control of coenzyme A levels in *Escherichia coli*. J. Bacteriol. 185, 3410– 3415.
- [56] Dosselaere, F. and Vanderleyden, J. (2001) A metabolic node in action: chorismate-utilizing enzymes in microorganisms. Crit. Rev. Microbiol. 27, 75–131.
- [57] Rolain, J.M., Stuhl, L., Maurin, M. and Raoult, D. (2002) Evaluation of antibiotic susceptibilities of three rickettsial species including *Rickettsia felis* by a quantitative PCR DNA assay. Antimicrob. Agents Chemother. 46, 2747–2751.
- [58] Eisenberg, M.A. (1985) Regulation of the biotin operon in *E. coli*. Ann. N.Y. Acad. Sci. 447, 335–349.
- [59] Rodionov, D.A., Mironov, A.A. and Gelfand, M.S. (2002) Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. Genome Res. 12, 1507–1516.

- [60] Amano, K.I., Williams, J.C. and Dasch, G.A. (1998) Structural properties of lipopolysaccharides from *Rickettsia typhi* and *Rickettsia prowazekii* and their chemical similarity to the lipopolysaccharide from *Proteus vulgaris OX19* used in the Weil–Felix test. Infect. Immun. 66, 923–926.
- [61] Lin, M. and Rikihisa, Y. (2003) *Ehrlichia chaffeensis* and *Anaplasma phagocytophilum* lack genes for Lipid A biosynthesis and incorporate cholesterol for their survival. Infect. Immun. 71, 5324–5331.
- [62] Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407, 81–86.
- [63] Gilmore Jr., R.D. and Hackstadt, T. (1991) DNA polymorphism in the conserved 190 kDa antigen gene repeat region among spotted fever group *Rickettsiae*. Biochim. Biophys. Acta 1097, 77–80.
- [64] Gilmore Jr., R.D. (1993) Comparison of the rompA gene repeat regions of *Rickettsiae* reveals species-specific arrangements of individual repeating units. Gene 125, 97–102.
- [65] Xu, W. and Raoult, D. (1997) Distribution of immunogenic epitopes on the two major immunodominant proteins (rOmpA and rOmpB) of *Rickettsia conorii* among the other rickettsiae of the spotted fever group. Clin. Diagn. Lab. Immunol. 4, 753–763.
- [66] Roux, V., Fournier, P.E. and Raoult, D. (1996) Differentiation of spotted fever group rickettsiae by sequencing and analysis of restriction fragment length polymorphism of PCR-amplified DNA of the gene encoding the protein rOmpA. J. Clin. Microbiol. 34, 2058–2065.
- [67] Li, H. and Walker, D.H. (1998) rOmpA is a critical protein for the adhesion of *Rickettsia rickettsii* to host cells. Microb. Pathog. 24, 289–298.
- [68] Van Der Laan, M., Urbanus, M.L., Ten Hagen-Jongman, C.M., Nouwen, N., Oudega, B., Harms, N., Driessen, A.J. and Luirinck, J. (2003) A conserved function of YidC in the biogenesis of respiratory chain complexes. Proc. Natl. Acad. Sci. USA 100, 5801–5806.
- [69] Joly, J.C. and Wickner, W. (1993) The SecA and SecY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids. EMBO J. 12, 255– 263.
- [70] Higgins, C.F. (1992) ABC transporters: from microorganisms to man. Annu. Rev. Cell. Biol. 8, 67–113.
- [71] Page, M.D., Pearce, D.A., Norris, H.A. and Ferguson, S.J. (1997) The *Paracoccus denitrificans* ccmA, B and C genes: cloning and sequencing, and analysis of the potential of their products to form a haem or apo-c-type cytochrome transporter. Microbiology 143, 563–576.
- [72] Schulz, H., Fabianek, R.A., Pellicioli, E.C., Hennecke, H. and Thony-Meyer, L. (1999) Heme transfer to the heme chaperone CcmE during cytochrome *c* maturation requires the CcmC protein, which may function independently of the ABCtransporter CcmAB. Proc. Natl. Acad. Sci. USA 96, 6462– 6467.
- [73] Putman, M., van Veen, H.W. and Konings, W.N. (2000) Molecular properties of bacterial multidrug transporters. Microbiol. Mol. Biol. Rev. 64, 672–693.
- [74] Fekkes, P. and Driessen, A.J. (1999) Protein targeting to the bacterial cytoplasmic membrane. Microbiol. Mol. Biol. Rev. 63, 161–173.
- [75] Mori, H. and Ito, K. (2001) The Sec protein-translocation pathway. Trends Microbiol. 1, 494–500.
- [76] Driessen, A.J., Fekkes, P. and van der Wolk, J.P. (1998) The Sec system. Curr. Opin. Microbiol. 1, 216–222.
- [77] Berks, B.C., Sargent, F. and Palmer, T. (2000) The Tat protein export pathway. Mol. Microbiol. 35, 260–274.

- [78] Powers, T. and Walter, P. (1997) Co-translational protein targeting catalyzed by the *Escherichia coli* signal recognition particle and its receptor. EMBO J. 16, 4880–4886.
- [79] Christie, P.J. and Vogel, J.P. (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol. 8, 354–360.
- [80] Ohashi, N., Zhi, N., Lin, Q. and Rikihisa, Y. (2002) Characterization and transcriptional analysis of gene clusters for a type IV secretion machinery in human granulocytic and monocytic ehrlichiosis agents. Infect. Immun. 70, 2128–2138.
- [81] Felek, S., Huang, H. and Rikihisa, Y. (2003) Sequence and expression analysis of virB9 of the type IV secretion system of *Ehrlichia canis* strains in ticks, dogs, and cultured cells. Infect. Immun. 71, 6063–6067.
- [82] Paetzel, M., Dalbey, R.E. and Strynadka, N.C.J. (2000) The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target. Pharmacol. Ther. 87, 27–49.
- [83] Rahman, M.S., Simser, J.A., Macaluso, K.R. and Azad, A.F. (2003) Molecular and functional analysis of the *lepB* gene, encoding a Type I signal peptidase from *Rickettsia rickettsii* and *Rickettsia typhi*. J. Bacteriol. 185, 4578–4584.
- [84] Wood, J.M. (1999) Osmosensing by bacteria: signals and membrane-based sensors. Microbiol. Mol. Biol. Rev. 63, 230– 262.
- [85] Tjaden, J., Winkler, H.H., Schwoppe, C., Van Der Laan, M., Mohlmann, T. and Neuhaus, H.E. (1999) Two nucleotide transport proteins in *Chlamydia trachomatis*, one for net nucleoside triphosphate uptake and the other for transport of energy. J. Bacteriol. 181, 1196–1202.
- [86] Dunbar, S.A. and Winkler, H.H. (1997) Increased and controlled expression of the *Rickettsia prowazekii* ATP/ADP translocase and analysis of cysteine-less mutant translocase. Microbiology 143, 3661–3669.
- [87] Alexeyev, M.F. and Winkler, H.H. (1999) Membrane topology of the *Rickettsia prowazekii* ATP/ADP translocase revealed by novel dual pho-lac reporters. J. Mol. Biol. 285, 1503–1513.
- [88] Walker, T.S. and Winkler, H.H. (1978) Penetration of cultured mouse fibroblasts (L cells) by *Rickettsia prowazekii*. Infect. Immun. 22, 200–208.
- [89] Uchiyama, T. (2003) Adherence to and invasion of Vero cells by recombinant *Escherichia coli* expressing the outer membrane protein rOmpB of *Rickettsia japonica*. Ann. N.Y. Acad. Sci. 990, 585–590.
- [90] Diaz-Montero, C.M., Feng, H.M., Crocquet-Valdes, P.A. and Walker, D.H. (2001) Identification of protective components of two major outer membrane proteins of spotted fever group *Rickettsiae*. Am. J. Trop. Med. Hyg. 65, 371–378.
- [91] Feng, H.M., Whitworth, T., Olano, J.P., Popov, L. and Walker, D.H. (2004) Fc-dependent polyclonal antibodies and antibodies to outer membrane proteins A and B, but not to lipopolysaccharide, protect SCID mice against fatal *Rickettsia conorii* infection. Infect. Immun. 72, 2222–2228.
- [92] Walker, T.S. (1984) Rickettsial interactions with human endothelial cells in vitro: adherence and entry. Infect. Immun. 44, 205– 210.
- [93] Teysseire, N., Boudier, J.A. and Raoult, D. (1995) *Rickettsia conorii* entry into Vero cells. Infect. Immun. 63, 366–374.
- [94] Silverman, D.J., Santucci, L.A., Meyers, N. and Sekeyova, Z. (1992) Penetration of host cells by *Rickettsia rickettsii* appears to be mediated by a phospholipase of rickettsial origin. Infect. Immun. 60, 2733–2740.

- [95] Ramm, L.E. and Winkler, H.H. (1973) Rickettsial hemolysis: effect of metabolic inhibitors upon hemolysis and adsorption. J. Bacteriol. 7, 550–555.
- [96] Ramm, L.E. and Winkler, H.H. (1976) Identification of cholesterol in the receptor site for rickettsiae on sheep erythrocyte membranes. Infect. Immun. 13, 120–126.
- [97] Ojcius, D.M., Thibon, M., Mounier, C. and Dautry-Varsat, A. (1995) pH and calcium dependence of hemolysis due to *Rickettsia prowazekii*: comparison with phospholipase activity. Infect. Immun. 63, 3069–3072.
- [98] Gouin, E., Gantelet, H., Egile, C., Lasa, I., Ohayon, H., Villiers, V., Gounon, P., Sansonetti, P.J. and Cossart, P. (1999) A comparative study of the actin-based motilities of the pathogenic bacteria *Listeria monocytogenes*, *Shigella flexneri* and *Rickettsia conorii*. J. Cell Sci. 112, 1697–1708.
- [99] Heinzen, R.A. (2003) Rickettsial actin-based motility: behavior and involvement of cytoskeletal regulators. Ann. N.Y. Acad. Sci. 990, 535–547.
- [100] Finlay, B.B. and Falkow, S. (1997) Common themes in microbial pathogenicity revisited. Microbiol. Mol. Biol. Rev. 61, 136–169.
- [101] Hacker, J. and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54, 641–679.
- [102] Radulovic, S., Troyer, J.M., Beier, M.S., Lau, A.O.T. and Azad, A.F. (1999) Identification and molecular analysis of the gene encoding *Rickettsia typhi* hemolysin. Infect. Immun. 67, 6104– 6108.
- [103] Gaywee, J., Sacci Jr., J.B., Radulovic, S., Beier, M.S. and Azad, A.F. (2003) Subcellular localization of rickettsial invasion protein InvA. Am. J. Trop. Med. Hyg. 68, 92–96.
- [104] Renesto, P., Dehoux, P., Gouin, E., Touqui, L., Cossart, P. and Raoult, D. (2003) Identification and characterization of a phospholipase D-superfamily gene in rickettsiae. J. Infect. Dis. 188, 1276–1283.
- [105] Gouin, E., Egile, C., Dehoux, P., Villiers, V., Adams, J., Gertler, F., Li, R. and Cossart, P. (2004) The RickA protein of *Rickettsia conorii* activates the Arp2/3 complex. Nature 427, 457–461.
- [106] Cox, M.M. (1998) A broadening view of recombinational DNA repair in bacteria. Genes Cells 3, 65–78.
- [107] Eisen, J.A. and Hanawalt, P.C. (1999) A phylogenomic study of DNA repair genes, proteins, and processes. Mutat. Res. 435, 171–213.
- [108] Cabrera, M., Nghiem, Y. and Miller, J.H. (1988) mutM, a second mutator locus in *Escherichia coli* that generates G.C–T.A transversions. J. Bacteriol. 170, 5405–5407.
- [109] Dunkin, S.M. and Wood, D.O. (1994) Isolation and characterization of the *Rickettsia prowazekii recA* gene. J. Bacteriol. 176, 1777–1781.
- [110] Csonka, L.N. and Hanson, A.D. (1991) Prokaryotic osmoregulation: genetics and physiology. Annu. Rev. Microbiol. 45, 569–606.
- [111] Chatterji, D. and Ojha, A.K. (2001) Revisiting the stringent response, ppGpp and starvation signaling. Curr. Opin. Microbiol. 4, 160–165.
- [112] Claverie, J.M. and Ogata, H. (2003) The insertion of palindromic repeats in the evolution of proteins. Trends Biochem. Sci. 28, 75–80.
- [113] Renesto, P. and Raoult, D. (2003) From genes to proteins: in vitro expression of rickettsial proteins. Ann. N.Y. Acad. Sci. 990, 642–652.
- [114] Koonin, E.V. (2003) Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat. Rev. Microbiol. 1, 127–136.