universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

„Host adaptation and genome dynamics during experimental evolution of chlamydial symbionts“

verfasst von / submitted by

Paul Herrera BSc (Hons), MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 437
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Biologie / Biology
Betreut von / Supervisor:	Univ.-Prof. Dr. Matthias Horn

Acknowledgements

First of all I must say that it really feels awesome to actually be at this stage of my PhD journey, right at the very end. And even though finishing the doctoral studies is obviously the goal of anyone who starts a PhD, the trajectory from starting line to the final ribbon is by no means straight. I have many people to thank who helped me in different ways throughout this wonderful adventure in Vienna.

I would like to thank my advisor Matthias Horn who offered me this position, even though I did not have a strong microbiology background. He gave me the opportunity to develop my skills and work independently while also engaging in numerous intriguing discussions on our beloved bacteria. I am also very grateful to all the members of the symbiosis group at DoME who were so helpful during the PhD, be it in the lab or when I struggled with what should have been a simple line of code! I also have numerous fond memories of us laughing in the oval office, having one too many beers at the heuriger, and even enjoying some hilarious nights out at conferences. To Allen, Angelika, Astrid, Cecilia, Daryl, Florian, Jasmin, Lena, Lisa, Martin, Nadja, Patrick, Stephan, Tamara, Thomas and Vincent – it was a pleasure being your colleague and, more importantly, your friend.

I want to thank everyone at DoME for all the stimulating conversations, and anyone who offered me assistance, advice or words of wisdom when I needed them most, especially my fellow graduate students. I feel like reaching this stage has been a joint effort with so many of you, since we passed through so many similar emotions, ups and downs, and we collectively shared our experiences and helped each other throughout. I look forward to seeing you all cross the finish line as well.

And finally, I would like to thank my close friends and family, in particular my parents, who have always supported me in whatever I chose to do and believed I would succeed, even when I perhaps did not.

Thank you all for sharing this experience with me. I am excited to see where my path will lead to next.

Table of Contents

Chapter I

Introduction

Introduction

Prokaryotes are everywhere. Any environment, as extreme as it may be, is replete with bacteria - from deep-sea hydrothermal vents to the freezing expanses of Antarctic soil. Even our bodies harbour more bacterial cells than human cells (Sender et al. 2016)! As the American writer Stewart Brand aptly put it: “If you don't like bacteria, you're on the wrong planet.” It should therefore come as no surprise that bacteria have also exploited eukaryotic cells, and nature is plentiful with examples of these host-associated bacteria. These microbes have major impacts on the biology, ecology and evolution of their hosts, affecting processes and functions including nutrition (Bäckhed et al. 2005) and development (Braendle et al. 2003), as well as speciation (Hurst and Werren 2001) and immunity (Macdonald and Monteleone 2005). Whereas pathogens tend to have major negative impacts on the hosts that they interact with, intracellular bacteria may become persistent and such sustained infections result in symbioses.

Obligate symbionts and their hosts often share a mutualistic interaction, and they descend from ancient and specialised associations. As a result of the very intimate relationship between host and endosymbiont, it is generally extremely difficult to culture these microorganisms, making them difficult to study. Thus, our knowledge about these elusive microbes lags behind. What we do know about genome evolution of obligate intracellular bacteria is limited to very few groups of organisms. Comparative genome analysis of the mutualistic pea aphid symbiont, *Buchnera aphidicola*, has provided us with major insights into the process of genome size reduction, genome dynamics of small genomes as well as adaptation to an obligate mutualistic relationship with a eukaryotic host (Pérez-Brocal et al. 2006; Moran et al. 2009).

This thesis focuses on a fascinating group of bacteria, the phylum *Chlamydiae*, which very likely represent the most ancient group of obligate intracellular bacteria. Comparative and phylogenetic genome sequence analysis have suggested that the last common ancestor of all extant members of this group already inhabited a eukaryotic host cell several hundreds of millions of years ago. Members of the chlamydiae possess small, reduced genomes and lack essential biosynthetic

pathways, relying on numerous metabolites from their eukaryotic hosts (Omsland et al. 2014). All *Chlamydiae* share a characteristic biphasic developmental cycle (fig. 1) comprising two distinct morphological and physiological stages: the elementary body (EB) survives the extracellular environment and infects new host cells, whereas the reticulate body (RB) replicates inside a host-derived vacuole (Abdelrahman and Belland 2005; Horn 2008). This developmental cycle is most certainly a unique feature among obligate intracellular bacteria.

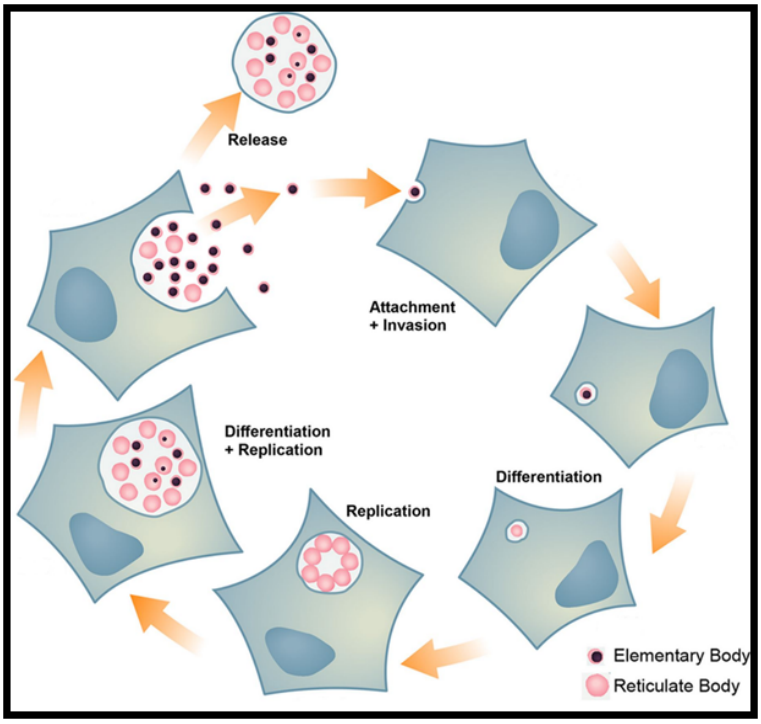

FIG. 1. Chlamydial biphasic developmental cycle (Omsland et al. 2014). An extracellular body (EB) is initially taken up by a host cell. This may then either form a single-cell inclusion or else multiple EBs may reside in one inclusion. Here the EB differentiates into the fully metabolic reticulate body (RB), which in turn replicates. Eventually, RBs convert back to EBs and are released from the host cell. This may result in host cell lysis but when this does not happen, the host and symbiont are able to co-evolve together.

The chlamydiae are especially relevant in today's world as they are very successful bacterial pathogens of humans. *Chlamydia trachomatis*, the first identified member of the group, comprises more than fifteen serologically defined variants and causes over 100 million infections worldwide every year (Wright et al. 2008; Roulis et al. 2013; Dal Conte et al. 2014). It is the main cause of preventable blindness in the developing world and the largest cause of bacterial sexually transmitted infections in the developed world. *C. pneumoniae* is a causative agent of pneumonia in humans

but it also infects numerous mammals, reptiles, and amphibians (Horn 2008; Taylor-Brown et al. 2015). In fact, other members of the chlamydiae can also infect a large variety of animals, including fish, birds and mammals, causing numerous clinically important diseases in economically significant livestock species (cattle, goats, sheep, and pigs) (DeGraves et al. 2003; Polkinghorne et al. 2009; Schautteet and Vanrompay 2011). However, in 1997 chlamydia-related organisms were unexpectedly discovered within amoeba cells isolated from human nasal mucosa (Amann et al. 1997), and nowadays numerous families of chlamydiae have been described (fig. 2) and are collectively referred to as the environmental chlamydiae (Horn 2008; Collingro et al. 2020).

These particular members of the chlamydiae are found ubiquitously in the environment and the majority of them are considered to be symbionts of free-living amoebae, with a limited ability to multiply in nonprotozoan host cells (Maurin et al. 2002; Greub et al. 2003; Collingro et al. 2005; Casson et al. 2006; Kebbi-Beghdadi et al. 2011; Sixt et al. 2012). Such protozoa have actually played a crucial role in the evolution of intracellular bacterial pathogens, which has been referred to as the ‘protozoan gym’ hypothesis (Harb et al. 2000). Free-living amoebae that harbour these symbionts are present in marine and freshwater habitats, and they are the main predators of bacteria in soil. Through phagocytosis, nutrients are released and taken up by plants, and it has been shown that this phenomenon, termed the soil microbial loop, is critical in driving plant productivity. Despite this major importance of protists in the soil community, very little is known on the effects that intracellular symbionts, which are consistently detected in (25% to 100% of) *Acanthamoeba* isolates, may incur on the activity of these protists.

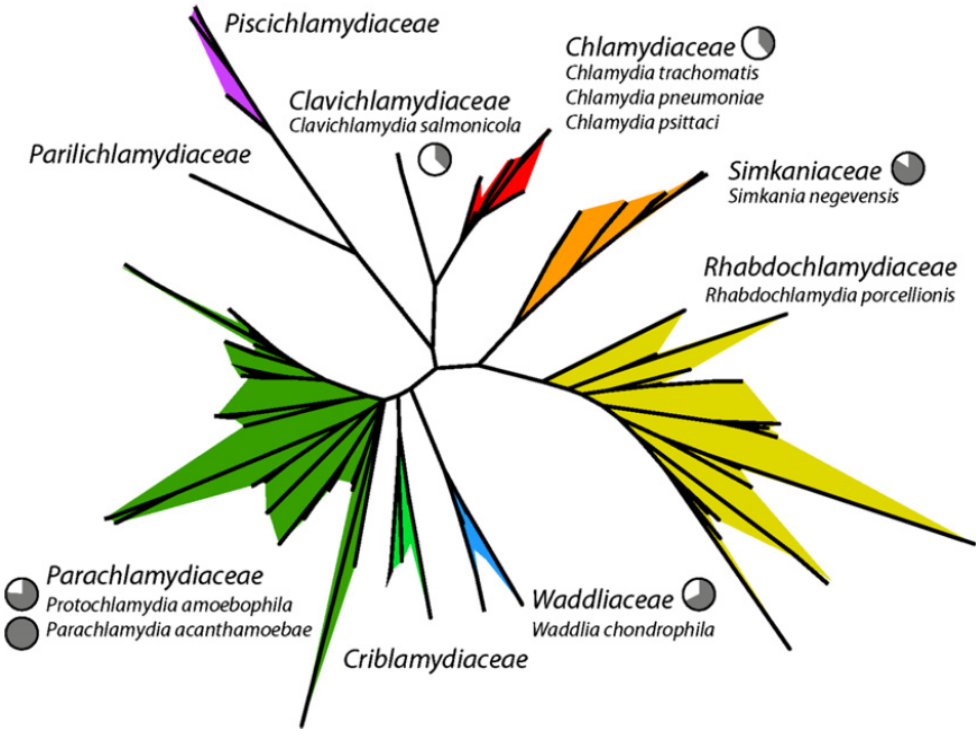

FIG. 2. The phylogenetic tree of the *Chlamydiae* (Omsland et al. 2014). The relationship between the pathogenic chlamydiae, the *Chlamydiaceae*, with the other families of environmental chlamydiae is shown. Representative species are depicted, with their genome sizes relative to the 3.1 Mb genome of *Parachlamydia acanthamoebae* indicated as pie charts.

Unlike well characterized obligate intracellular bacteria such as *B. aphidicola* and *Wolbachia* spp., the *Chlamydiae* have only been poorly studied with respect to genome evolution and dynamics. However, I believe that they provide an exciting prospect toward a better understanding of major drivers in the evolution of host-associated microbes due to their unique features and significance in our society. Experimental evolution—the study of evolutionary processes in response to imposed controlled conditions in real time—using microbes has become increasingly popular over the last few decades as they are ideal subjects (Elena and Lenski 2003; Buckling et al. 2009; Brockhurst et al. 2010; Dettman et al. 2012; Cvijović et al. 2018). Long-term experiments with microorganisms are brilliant as they can be used as a low-complexity model system in order to test different evolutionary theories. One of the finest examples to date of such an experiment is that initiated by Richard Lenski in 1988 using 12 serially propagated *Escherichia coli* populations in minimal medium (Lenski et al. 1991), which is still ongoing. Over 80 publications on this long-term experiment alone have yielded invaluable insights into the evolutionary

dynamics of adaptation (Barrick et al. 2009; Wiser et al. 2013; Good et al. 2017), fitness trajectories (Lenski et al. 2015), the evolution of mutation rates (Sniegowski et al. 1997; Arjan et al. 1999) and so many other fine details in the evolutionary process. Apart from studying evolution, experimental evolution using microbes can be applied to other fields as well such as cancer research (Nowell 1976; Sprouffske et al. 2012) and antibiotic resistance (Stone et al. 2016; Yen and Papin 2017).

In the past, such studies were limited to cultivable, easily accessible and genetically tractable microbes (Elena and Lenski 2003; Buckling et al. 2009), but the advent of highly parallel sequencing methods now allows us to explore the more elusive host-associated bacteria. We can finally shed some light on how such intricate relationships have stood the test of time for so long, with one area to explore involving the character of the symbioses. The relationship between the partners in microbial symbioses has been described as existing along a parasitism-mutualism continuum (Ewald 1987), whose dynamics depend on diverse genotypic and environmental factors (Michalakis et al. 1992; Lipsitch et al. 1996; Yamamura 1996; Hoeksema et al. 2010). Transmission mode is of particular relevance and is an important factor shaping the ecology and evolution of numerous symbiotic associations (Yamamura 1993; Moran et al. 2008; Bright and Bulgheresi 2010). In vertical transmission, symbionts are inherited from the host parent to the offspring, aligning the evolutionary interest of host and symbionts. This selects for benign, mutualistic symbionts (Axelrod and Hamilton 1981; Bull and Rice 1991). By contrast, horizontally transmitted symbionts infect new host lineages and this promotes the evolution of virulence (Fine 1975; Ewald 1983; Bull 1994). Numerous experimental studies have validated these arguments (Bull et al. 1991; Messenger et al. 1999; Stewart et al. 2005; Sachs and Wilcox 2006) but, interestingly, we know close to nothing about molecular adaptations leading to these shifts in the parasitism-mutualism continuum. Another important aspect relating to the chlamydiae is how these bacteria made the leap to higher eukaryotes. Temperature adaptation is essential for many microbes, particularly pathogens that invade warm-blooded animals. Such adaptations regularly involve the upregulation of temperature-regulated genes. Often this needs global gene regulation involving a two-component signal transduction system composed of a sensor-kinase and a response regulator (Konkel and Tilly 2000).

The central theme of this thesis involves the application of experimental evolution approaches to unravel aspects of the chlamydia-amoeba symbiosis. I aim to understand the genomic and molecular basis of the intracellular lifestyle of two species of environmental chlamydiae (family *Parachlamydiaceae*) with their *Acanthamoeba* host with respect to aspects of host interaction and adaptation, as well as the character of the symbioses. The third chapter dives right into this latter aspect, where we make use of *Parachlamydia acanthamoebae* to understand the position of the symbiosis along the parasitism-mutualism continuum. We designed two regimes by experimentally manipulating the transmission of symbionts among host generations to predominantly select for either of the two transmission modes (horizontal or vertical). The fourth chapter makes use of *Protochlamydia amoebophila*, where we set up an evolution experiment for a number of years, at 20 °C and at an elevated, more stressful 30 °C. Here I really am interested in tracking genomic evolutionary dynamics as well as observing any evolved phenotypes. The fifth chapter goes one step deeper into this evolution experiment where we use transcriptomics to explore how the symbiont adapts to sustained elevated temperature for long periods of time and to identify those pathways that were most affected by the high temperature.

References

Abdelrahman YM, Belland RJ. 2005. The chlamydial developmental cycle. *FEMS Microbiol. Rev.* 29:949–959.

Amann R, Springer N, Schönhuber W, Ludwig W, Schmid EN, Müller KD, Michel R. 1997. Obligate intracellular bacterial parasites of acanthamoebae related to *Chlamydia* spp. *Appl. Environ. Microbiol.* 63:115–121.

Arjan JA, Visser M, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science* 283:404–406.

Axelrod R, Hamilton WD. 1981. The evolution of cooperation. *Science* 211:1390–1396.

Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. 2005. Host-bacterial mutualism in the human intestine. *Science* 307:1915–1920.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Braendle C, Miura T, Bickel R, Shingleton AW, Kambhampati S, Stern DL. 2003. Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. *PLoS Biol.* 1:E21.

Bright M, Bulgheresi S. 2010. A complex journey: transmission of microbial symbionts. *Nat. Rev. Microbiol.* 8:218–230.

Brockhurst MA, Colegrave N, Rozen DE. 2010. Next-generation sequencing as a tool to study microbial evolution. *Mol. Ecol.* 20:972–980.

Buckling A, Craig Maclean R, Brockhurst MA, Colegrave N. 2009. The Beagle in a bottle. *Nature* 457:824–829.

Bull JJ. 1994. Perspective: Virulence. *Evolution* [Internet] 48:1423. Available from: http://dx.doi.org/10.2307/2410237

Bull JJ, Molineux IJ, Rice WR. 1991. Selection of Benevolence in a Host-Parasite System. *Evolution* [Internet] 45:875. Available from: http://dx.doi.org/10.2307/2409695

Bull JJ, Rice WR. 1991. Distinguishing mechanisms for the evolution of co-operation. *J. Theor. Biol.* 149:63–74.

Casson N, Medico N, Bille J, Greub G. 2006. *Parachlamydia acanthamoebae* enters and multiplies within pneumocytes and lung fibroblasts. *Microbes Infect.* 8:1294–1300.

Collingro A, Köstlbacher S, Horn M. 2020. *Chlamydiae* in the Environment. *Trends Microbiol.* [Internet]. Available from: http://dx.doi.org/10.1016/j.tim.2020.05.020

Collingro A, Poppert S, Heinz E, Schmitz-Esser S, Essig A, Schweikert M, Wagner M, Horn M. 2005. Recovery of an environmental *Chlamydia* strain from activated sludge by co-cultivation with *Acanthamoeba* sp. *Microbiology* 151:301–309.

Cvijović I, Nguyen Ba AN, Desai MM. 2018. Experimental Studies of Evolutionary Dynamics in Microbes. *Trends Genet.* 34:693–703.

Dal Conte I, Mistrangelo M, Cariti C, Chiriotto M, Lucchini A, Vigna M, Morino M, Di Perri G. 2014. Lymphogranuloma venereum: an old, forgotten re-emerging systemic disease. *Panminerva Med.* 56:73–83.

DeGraves FJ, Gao D, Hehnen H-R, Schlapp T, Kaltenboeck B. 2003. Quantitative Detection of *Chlamydia psittaci* and *C. pecorum* by High-Sensitivity Real-Time PCR Reveals High Prevalence of Vaginal Infection in Cattle. *J. Clin. Microbiol.* [Internet] 41:1726–1729. Available from: http://dx.doi.org/10.1128/jcm.41.4.1726-1729.2003

Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF, Kassen R. 2012. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Mol. Ecol.* 21:2058–2077.

Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4:457–469.

Ewald PW. 1983. Host-Parasite Relations, Vectors, and the Evolution of Disease Severity. *Annu. Rev. Ecol. Syst.* [Internet] 14:465–485. Available from: http://dx.doi.org/10.1146/annurev.es.14.110183.002341

Ewald PW. 1987. Transmission modes and evolution of the parasitism-mutualism continuum. *Ann. N. Y. Acad. Sci.* 503:295–306.

Fine PEM. 1975. Vectors and vertical transmission: an epidemiologic perspective. *Ann. NY Acad. Sci.* [Internet] 266:173–194. Available from: http://dx.doi.org/10.1111/j.1749-6632.1975.tb35099.x

Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature* 551:45–50.

Greub G, Mege J-L, Raoult D. 2003. *Parachlamydia acanthamoebae* enters and multiplies within human macrophages and induces their apoptosis [corrected]. *Infect. Immun.* 71:5979–5985.

Harb OS, Gao L-Y, Kwaik YA. 2000. From protozoa to mammalian cells: a new paradigm in the life cycle of intracellular bacterial pathogens. Minireview. *Environ. Microbiol.* [Internet] 2:251–265. Available from: http://dx.doi.org/10.1046/j.1462-2920.2000.00112.x

Hoeksema JD, Chaudhary VB, Gehring CA, Johnson NC, Karst J, Koide RT, Pringle A, Zabinski C, Bever JD, Moore JC, et al. 2010. A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. *Ecol. Lett.* 13:394–407.

Horn M. 2008. *Chlamydiae* as symbionts in eukaryotes. *Annu. Rev. Microbiol.* 62:113–131.

Hurst GD, Werren JH. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.* 2:597–606.

Kebbi-Beghdadi C, Batista C, Greub G. 2011. Permissivity of fish cell lines to three *Chlamydia*-related bacteria: *Waddlia chondrophila*, *Estrella lausannensis* and *Parachlamydia acanthamoeba*e. *FEMS Immunol. Med. Microbiol.* 63:339–345.

Konkel ME, Tilly K. 2000. Temperature-regulated expression of bacterial virulence genes. *Microbes Infect.* 2:157–166.

Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *Am. Nat.* 138:1315–1341.

Lenski RE, Wiser MJ, Ribeck N, Blount ZD, Nahum JR, Morris JJ, Zaman L, Turner CB, Wade BD, Maddamsetti R, et al. 2015. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. *Proc. Biol. Sci.* 282:20152292.

Lipsitch M, Siller S, Nowak MA. 1996. The evolution of virulence in pathogens with vertical and horizontal transmission. *Evolution* 50:1729–1741.

Maurin M, Bryskier A, Raoult D. 2002. Antibiotic susceptibilities of *Parachlamydia acanthamoeba* in amoebae. *Antimicrob. Agents Chemother.* 46:3065–3067.

Messenger SL, Molineux IJ, Bull JJ. 1999. Virulence evolution in a virus obeys a trade off. *P. Roy. Soc. Lond. B Bio.* [Internet] 266:397–404. Available from: http://dx.doi.org/10.1098/rspb.1999.0651

Michalakis Y, Olivieri I, Renaud F, Raymond M. 1992. Pleiotropic action of parasites: How to be good for the host. *Trends Ecol. Evol.* 7:59–62.

Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42:165–190.

Moran NA, McLaughlin HJ, Sorek R. 2009. The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. *Science* 323:379–382.

Nowell P. 1976. The clonal evolution of tumor cell populations. *Science* [Internet] 194:23–28. Available from: http://dx.doi.org/10.1126/science.959840

Omsland A, Sixt BS, Horn M, Hackstadt T. 2014. Chlamydial metabolism revisited: interspecies metabolic variability and developmental stage-specific physiologic activities. *FEMS Microbiol. Rev.* 38:779–801.

Pérez-Brocal V, Gil R, Ramos S, Lamelas A, Postigo M, Michelena JM, Silva FJ, Moya A, Latorre A. 2006. A small microbial genome: the end of a long symbiotic relationship? *Science* 314:312–313.

Polkinghorne A, Borel N, Becker A, Lu ZH, Zimmermann DR, Brugnera E, Pospischil A, Vaughan L. 2009. Molecular evidence for chlamydial infections in the eyes of sheep. *Vet. Microbiol.* 135:142–146.

Roulis E, Polkinghorne A, Timms P. 2013. *Chlamydia pneumoniae*: modern insights into an ancient pathogen. *Trends Microbiol.* 21:120–128.

Sachs JL, Wilcox TP. 2006. A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc. Biol. Sci.* 273:425–429.

Schautteet K, Vanrompay D. 2011. *Chlamydiaceae* infections in pig. *Vet. Res.* [Internet] 42:29. Available from: http://dx.doi.org/10.1186/1297-9716-42-29

Sender R, Fuchs S, Milo R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* 14:e1002533.

Sixt BS, Hiess B, König L, Horn M. 2012. Lack of effective anti-apoptotic activities restricts growth of *Parachlamydiaceae* in insect cells. *PLoS One* 7:e29565.

Sniegowski PD, Gerrish PJ, Lenski RE. 1997. Evolution of high mutation rates in experimental populations of *E. coli*. *Nature* 387:703–705.

Sprouffske K, Merlo LMF, Gerrish PJ, Maley CC, Sniegowski PD. 2012. Cancer in Light of Experimental Evolution. *Curr. Biol.* [Internet] 22:R762–R771. Available from: http://dx.doi.org/10.1016/j.cub.2012.06.065

Stewart AD, Logsdon JM Jr, Kelley SE. 2005. An empirical study of the evolution of virulence under both horizontal and vertical transmission. *Evolution* 59:730–739.

Stone LK, Baym M, Lieberman TD, Chait R, Clardy J, Kishony R. 2016. Compounds that select against the tetracycline-resistance efflux pump. *Nat. Chem. Biol.* 12:902–904.

Taylor-Brown A, Vaughan L, Greub G, Timms P, Polkinghorne A. 2015. Twenty years of research into *Chlamydia*-like organisms: a revolution in our understanding of the biology and pathogenicity of members of the phylum *Chlamydiae*. *Pathog. Dis.* 73:1–15.

Wiser MJ, Ribeck N, Lenski RE. 2013. Long-term dynamics of adaptation in asexual populations. *Science* 342:1364–1367.

Wright HR, Turner A, Taylor HR. 2008. Trachoma. *Lancet* 371:1945–1954.

Yamamura N. 1993. Vertical Transmission and Evolution of Mutualism from Parasitism. *Theor. Popul. Biol.* [Internet] 44:95–109. Available from: http://dx.doi.org/10.1006/tpbi.1993.1020

Yamamura N. 1996. Evolution of mutualistic symbiosis: A differential equation model. *Res. Popul. Ecol.* [Internet] 38:211–218. Available from: http://dx.doi.org/10.1007/bf02515729

Yen P, Papin JA. 2017. History of antibiotic adaptation influences microbial evolutionary dynamics during subsequent treatment. *PLoS Biol.* 15:e2001586.

Chapter II

Overview of manuscripts

Chapter III

Manuscript title:

Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont

Author names:

Paul Herrera, Lisa Schuster, Cecilia Wentrup, Lena König, Thomas Kempinger, Hyunsoo Na, Jasmin Schwarz, Stephan Köstlbacher, Florian Wascher, Markus Zojer, Thomas Rattei, Matthias Horn

Reference:

Proceedings of the National Academy of Sciences of the United States of America (August 19, 2020). doi:10.1073/pnas.2005536117.

Author contributions

MH conceived the study. PH, LS, LK, CW and MH designed the experiments. LS, JS, FW and HN performed the evolution experiment. PH and TK performed the fitness assays. PH, CW, and JS performed the genome and transcriptome sequencing experiments. PH, LK, MZ and MH analysed fitness, genome and transcriptome data. SK and TR provided data analysis support. PH and MH wrote the manuscript. PH, LS, CW, LK, SK, MZ, TR, and MH edited the manuscript.

Chapter IV

Manuscript title:

Genome dynamics and temperature adaptation during experimental evolution of obligate intracellular bacteria

Author names:

Paul Herrera, Lisa Schuster, Markus Zojer, Hyunsoo Na, Jasmin Schwarz, Florian Wascher, Thomas Kempinger, Andreas Regner, Thomas Rattei, Matthias Horn

Reference:

This manuscript was recently submitted to *Molecular Biology and Evolution*.

Author contributions:

MH conceived the study. LS and MH designed the experiments. LS, JS, FW, HN, PH and TK performed the evolution experiment. PH, AR and TK performed the fitness assays. PH, LS and JS performed the genome sequencing experiments. PH, MZ, TK and MH analysed fitness and genome data. TR provided data analysis support. PH and MH wrote the manuscript. All authors edited and approved the final manuscript.

Chapter V

Manuscript title:

Survival at an elevated temperature strongly reduces gene expression dynamics in a chlamydial symbiont

Author names:

Paul Herrera, Cecilia Wentrup, Lena König, Stephan Köstlbacher, Thomas Rattei, Matthias Horn

Reference:

Draft manuscript

Author contributions:

MH conceived the study. PH, LK, CW and MH designed the experiments. PH and CW performed the RNA sequencing experiments. PH, LK and MH analysed the transcriptome data. SK and TR provided data analysis support. PH wrote the manuscript.

Chapter III

Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont

Title:

Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont

Authors:

Paul Herrera[a], Lisa Schuster[a], Cecilia Wentrup[a], Lena König[a], Thomas Kempinger[a], Hyunsoo Na[a,b], Jasmin Schwarz[a], Stephan Köstlbacher[a], Florian Wascher[a], Markus Zojer[a], Thomas Rattei[a], Matthias Horn[a]

Author affiliations:

[a] Centre for Microbiology and Environmental Systems Science, University of Vienna

[b] Current address: Department of Energy Joint Genome Institute, Walnut Creek, CA, USA

Corresponding author:

Univ.-Prof. Dr. Matthias Horn

University of Vienna

Department of Microbiology and Ecosystem Science

Althanstr. 14

A-1090 Wien

Austria

phone: +43 1 4277 91208

email: matthias.horn@univie.ac.at

Abstract

Symbiosis with microbes is a ubiquitous phenomenon with a massive impact on all living organisms, shaping the world around us today. Theoretical and experimental studies show that vertical transmission of symbionts leads to the evolution of mutualistic traits, whereas horizontal transmission facilitates the emergence of parasitic features. However, these studies focused on phenotypic data, and we know little about underlying molecular changes at the genomic level. Here, we combined an experimental evolution approach with infection assays, genome re-sequencing, and global gene expression analysis to study the effect of transmission mode on an obligate intracellular bacterial symbiont. We show that a dramatic shift in the frequency of genetic variants, coupled with major changes in gene expression, allow the symbiont to alter its position in the parasitism-mutualism continuum depending on the mode of between-host transmission. We found that increased parasitism in horizontally transmitted chlamydiae residing in amoebae was a result of processes occurring at the infectious stage of the symbiont's developmental cycle. Specifically, genes involved in energy production required for extracellular survival and the type III secretion system - the symbiont's primary virulence mechanism - were significantly upregulated. Our results identify the genomic and transcriptional dynamics sufficient to favor parasitic or mutualistic strategies.

Significance statement

Symbiotic relationships with microbes are ubiquitous among living beings and can be parasitic, such as in bacterial pathogens, or mutualistic, as in beneficial microbiomes. Among other factors, the outcome of microbe-host relationships is determined by the mode of symbiont transmission from host to host. Here we describe how bacterial symbionts increased in infectivity and virulence towards their amoeba host when transmission to a new host was essential for survival. The enhanced parasitism is a result of genomic changes and a pronounced switch of gene expression altering the symbionts' mechanisms for host interaction. Our study provides both a molecular explanation as well as a blueprint for how changes in gene expression are sufficient to confer enhanced parasitism in microbes.

Introduction

The relationship between the partners in microbial symbioses have been described as existing along a parasitism-mutualism continuum (1), and the dynamics of this continuum are dependent upon numerous genotypic and environmental factors (2–6). The transmission mode of viruses, bacteria, and eukaryotic parasites is a key factor shaping the ecology and evolution of symbiotic associations (7–10). In the case of vertical transmission (VT), symbionts are inherited from the host parent to offspring. By contrast, in horizontal transmission (HT), symbiont transmission is not linked to host reproduction, but instead symbionts infect new host lineages. The mode of symbiont transmission among host generations has important evolutionary ramifications: while VT aligns the evolutionary interest of host and symbionts, thus selecting for benign or even mutualistic symbionts, HT does not penalize parasitic strategies, because symbionts are associated with a new host individual every round (11–14). As a consequence, HT promotes the evolution of virulence (15–17), broadly defined as the decrease in host fitness as a result of infection, which can ultimately lead to host mortality (18–21). Consistent with this notion, experimental manipulation in the lab has demonstrated that propagating symbionts via either HT or VT will select for parasitic or mutualistic traits, respectively (19, 22–24). However, the molecular adaptations causing the observed shifts in the parasitism-mutualism continuum remain elusive.

Here, we take advantage of the bacterial symbiont *Parachlamydiaacanthamoebae* and its protist host *Acanthamoeba*sp.as a tractable model system. *Chlamydiae*, which represent the most ancient known group of obligate intracellular bacteria, all share a characteristic biphasic developmental cycle alternating between two distinct morphological and physiological stages: the elementary body (EB) survives (but cannot replicate) outside eukaryotic host cells and infects new hosts, whereas the reticulate body (RB) replicates inside the host cell within a host-derived vacuole (25–27). Chlamydiae are among the most successful bacterial pathogens of humans. Their environmental representatives (28) , such as the one used here, are ubiquitous in the environment and live primarily within unicellular protists like amoeba. During host cell division this symbiont can be vertically transmitted from parent to daughter cells (VT), but can also readily infect naive hosts (HT). This mixed

transmission mode (29) represents an ideal starting point for experimentally manipulating the transmission of symbionts among host generations (i.e. VT or HT), and in this way identify the molecular changes that arise in response to the different selection regimes.

Results and Discussion

Design of evolution experiment

In our VT regime, fully infected amoebae were allowed to replicate continuously and no naive hosts were added throughout the experiment (**Fig. 1**). EBs, the infectious chlamydial stage essential for horizontal transmission, were removed daily. This ensured that the extent of HT was reduced to a minimum and that the chlamydiae were predominantly transferred vertically from parent to daughter cells. In the HT regime EBs, which had been released from their amoeba host cells, were collected weekly and used to freshly infect naive amoebae. VT was minimized and HT was essential in this regime as all infected amoeba cells were removed at the end of the week, with only the infectious EBs being isolated and used for fresh infections. Both regimes were initiated from the same ancestral chlamydia and amoeba populations and maintained for a duration of 14 months, which corresponded to about 500 to 600 symbiont generations (*SI Appendix*, *SI Text*).

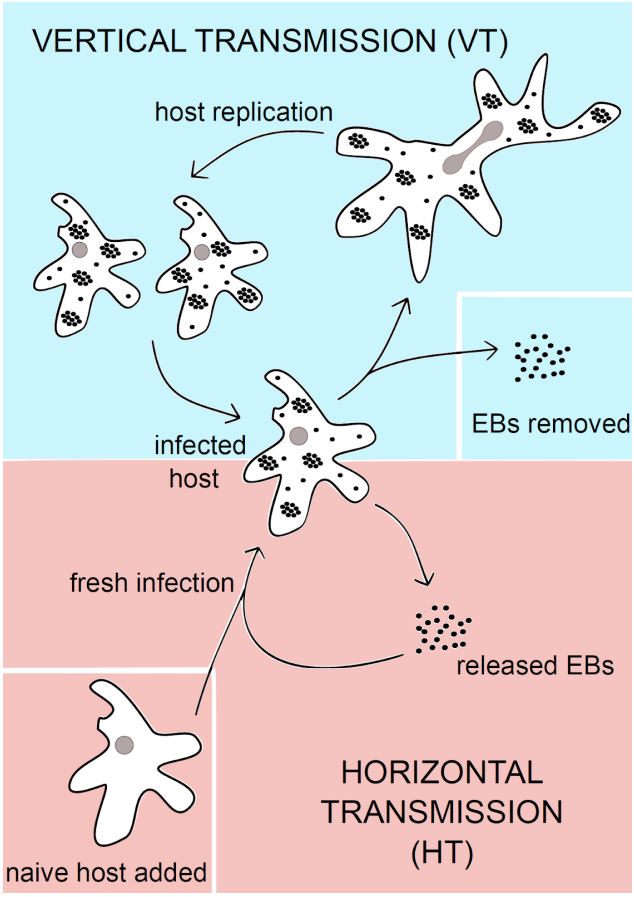

Fig. 1. The *Parachlamydia/Acanthamoeba* model and the design of the evolution experiment. Two separate selection regimes were established to select for HT and VT, both originating from the same initial ancestral amoeba-symbiont population. In VT (top), amoebae that were fully infected with the symbiont were maintained in liquid culture. Daily removal of the liquid medium was carried out to remove released EBs, minimizing the potential for HT. No new amoebae were added to the cultures, and the bacterial symbionts were passed on from parent cell to daughter cell. This was performed for 14 months, equivalent to 525 symbiont generations. In HT (bottom), EBs were allowed to accumulate over a week in the liquid medium, after which they were isolated and used to infect naive host cells. Infected amoebae were discarded at the end of the week. Fresh infections, by way of HT, were required for the symbionts to be maintained in the population. This was performed for 14 months, equivalent to 560 symbiont generations.

Fitness assays reveal a shift towards parasitism under HT conditions

To understand how symbiont and host fitness changed throughout the experiment, infection experiments were carried out to quantify growth parameters of both host and symbiont. Naive amoebae that were infected with symbionts isolated at the end of the experiment from the VT treatment displayed a net growth rate that was

statistically indistinguishable from the one of uninfected amoebae (**Fig. 2*A***). This observation indicated that vertically transmitted symbionts had minimal negative effects on the host. On the other hand, the net growth rate of naive amoebae freshly infected with symbionts isolated at the end of the experiment from the HT treatment was roughly half the net growth rate of both uninfected amoebae and amoebae from the VT treatment (**Fig. 2*A***). The observed pronounced decrease in host fitness when challenged with HT versus VT symbionts suggests a strong shift towards parasitism in horizontally transmitted symbionts.

The fitness of chlamydial symbionts during these infection experiments was quantified at two stages during the infection process. At 6 hours post infection (hpi) EBs have successfully entered the host cell, initiated differentiation to RBs and started the establishment of their intracellular niche. This time point thus determines the ability of the symbionts to infect the amoeba host at the start of the experiment. At 48 hpi, one chlamydial developmental cycle is complete, and the ability of the symbionts to rapidly multiply and subsequently be released back into the medium can be determined. Symbionts originating from the HT treatment were highly infectious from the start of the fitness assay: already 6 hpi, the symbiont numbers in amoebae infected with HT chlamydiae were double that of the VT treatment (**Fig. 2*B***). These elevated symbiont numbers were maintained until the completion of one developmental cycle (i.e. 48 hpi) (**Fig. 2*C***). Moreover, a significantly higher number of HT symbionts compared to VT symbionts was released per host cell at this time point (**Fig. 2*D***). These results demonstrate that chlamydiae isolated from the HT regime infected their host more efficiently than the ones isolated from the VT regime. This includes the initial uptake, formation and replication of RBs, and release of infectious EBs into the supernatant. The increased symbiont fitness reflects the observed reduction in host fitness (**Fig. 2*A***), together showing a pronounced shift towards parasitism under HT conditions. An increase in symbiont virulence and parasitism through HT meets theoretical expectations (15–17) . Although these predictions have been previously demonstrated for viruses and eukaryotic parasites, only very few studies exist so far that verified the link between transmission mode and the level of virulence in bacteria (30, 31).

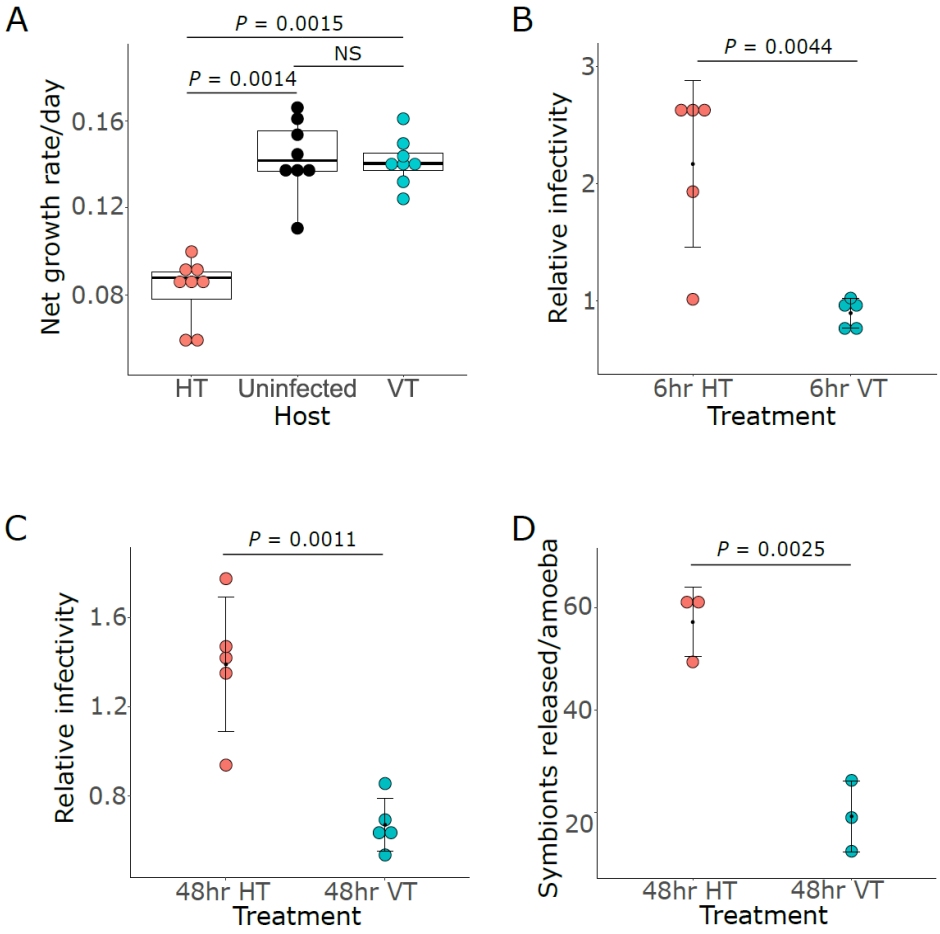

Fig. 2. Symbionts isolated from the HT regime are more virulent than symbionts from the VT regime. (A) Amoeba populations (n = 8) that were either fully infected with symbionts from the HT (red dots) or VT (blue dots) treatment, or uninfected (black dots), were seeded at a concentration of 2 x 10^4 cells per ml and incubated at 20 °C. The growth of the amoebae after 7 days was quantified and used to calculate the net growth rate per day. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. The rates were compared using a Kruskal-Wallis test (χ^2 (2) = 15.425, $P < 0.0005$) followed by Dunn's multiple comparison post-hoc test. NS, not significant. All growth rates are provided in Dataset S1. (B-D) Naive amoebae at a concentration of 10^5 cells per ml were infected with symbionts isolated from the HT (red dots) or VT (blue dots) treatment with a multiplicity of infection (MOI) of 5. The relative infectivity was determined by dividing the total symbiont number per ml (inside and outside host cells) of the particular treatment (n = 5) by that of the ancestral symbiont after 6 hpi (B) and 48 hpi (C). Data represent mean ± standard deviation of the mean. Numbers of bacteria were compared using an unpaired two-tailed Student's t test (B: $t(8) = 3.930$, $P < 0.005$; C: $t(8) = 4.989$, $P < 0.005$). All relative infectivity values are provided in Dataset S1. (D) Total numbers of symbionts released outside the host cells per amoeba for each treatment (n = 3) were compared using an unpaired two-tailed Student's t test ($t(4) = 6.792$,

$P < 0.005$). Data represent mean ± standard deviation of the mean. All individual values for total numbers of symbionts released are provided in Dataset S1.

Ancestral variants increased in frequency in the HT symbionts

Our phenotypic data suggests that the increase in symbiont virulence is a consequence of changes in the bacterial developmental cycle, emphasizing the efficiency of uptake and the generation and release of infectious bacteria into the environment. As a next step, we set out to investigate the molecular basis of these adaptations. The symbiont genome consists of a single circular chromosome (3,072,383 bp) with 2,618 genes (27). Pool sequencing of bacteria isolated from either treatment at the end of the experiment after 14 months allowed us to identify genomic changes that occurred in the symbiont populations during the experiment (**Fig. 3*A***). We identified 9 novel variants in the horizontal treatment and 24 novel variants in the vertical treatment, with many of them at low frequencies and none affecting genes involved in critical cellular functions (Dataset S2). Conversely, standing genetic variation in the initial ancestral population—the founder chlamydia and amoeba populations from which both regimes were initiated—appeared to be particularly important for the observed evolutionary changes. From a total of 1,436 ancestral variants, mainly single nucleotide polymorphisms (SNPs), spread around the genome, 1,161 variants (80.8% of all variants; Dataset S2) pronouncedly increased in frequency in the HT symbiont population as compared to the VT population, with a large majority of these variants reaching very high frequencies above 80% (**Fig. 3*A***). On the other hand, only 134 of these variants (9.3% of all variants; Dataset S2) increased in frequency in the VT population relative to the HT population. The strong shift in variant frequencies under HT conditions revealed a major change in the symbiont population concomitant with the observed phenotypic change in parasitism during our experiment.

In order to investigate this shift in variant frequencies, we focused on the set of genes that appeared most different between the two experimental conditions. For this, we focused on genes that included variants with a greater than 15% differential in frequency. These variants were distributed in a total of 446 genes accounting for 17% of the protein coding sequences, without apparent mutational hot spots. To

better understand how these genes are related to the observed phenotypes, we tested whether they are involved in specific cellular functions. Apart from considering the identity of the affected genes, the variants were divided into numerous sub-groups, based on whether they increased in frequency in the HT or the VT regime, as well as what type of variant it was (synonymous, nonsynonymous, or promoter region). Surprisingly, no statistically significant enrichments with respect to their predicted cellular functions were identified in any of these groups. Therefore, although genome re-sequencing of the symbionts revealed a major population shift and identified hundreds of genetic variants that increased in frequency under HT, the genome sequencing data alone does not provide any clues about the molecular basis underlying the observed phenotypes. We thus next asked whether insights into gene expression dynamics could help understand the difference in parasitism between HT and VT symbionts. Using bacteria isolated at the end of the experiment from either selection regime, we performed Illumina-based RNA sequencing (RNA-Seq) analysis on symbionts infecting naive amoebae isolated at three time points that mark crucial developmental events during infection: 2 hpi representing the start of infection and EB-to-RB transition; 24 hpi as the peak of replication and RB activity; and released EBs accumulated at 7 days post infection, representing the new generation of symbionts ready to infect new host cells.

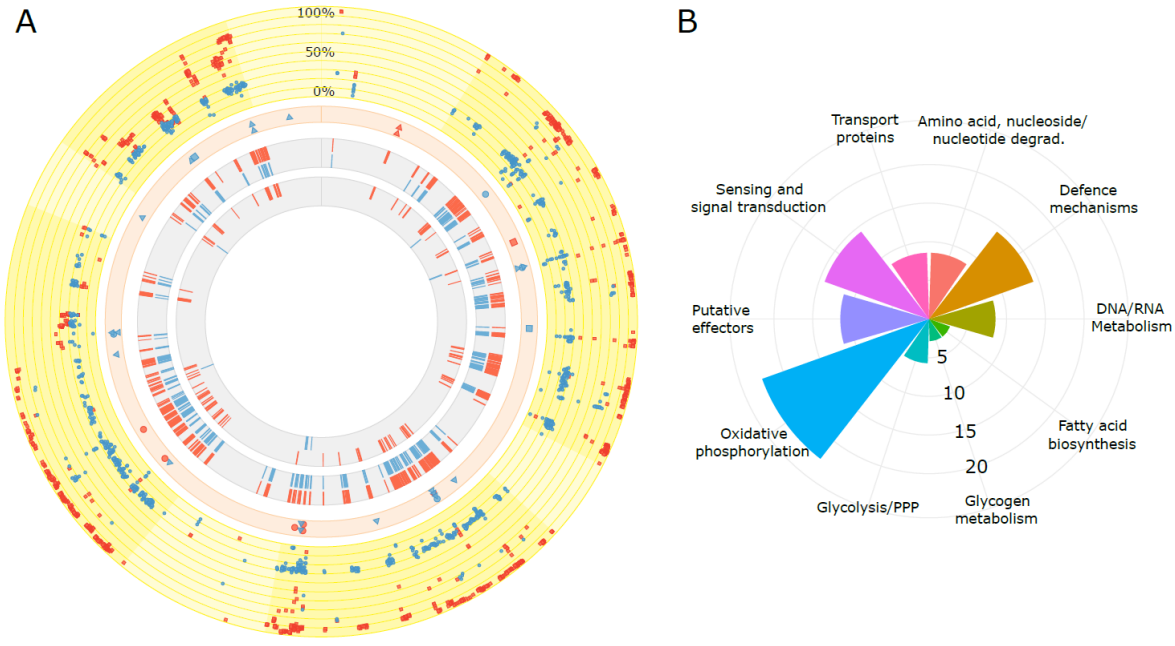

Fig. 3. A majority of genetic variants strongly increased in frequency in the symbionts from the HT regime, and were associated with differentially expressed genes. The circular plots represent the symbiont genome. (A) In all four rings, red indicates HT and blue indicates VT.

The outermost ring depicts a scatter plot for the entire symbiont genome with all the 1,436 ancestral genetic variants identified in the HT and VT symbiont populations. A large majority of the HT variants reach very high frequencies above 80%. One symbol represents a single variant. Lighter yellow sections represent conserved regions of the genome that possess a high proportion of housekeeping genes and are less affected by variants (i.e. genes involved in replication, translation, central carbon metabolism, energy production and conversion, metabolism and transport). The second ring shows the location of the 33 novel variants (9 = HT; 24 = VT) identified throughout the genome (circle = intergenic; triangle = genic; rectangle = potential promoter region). The third ring depicts those genes from either treatment (354 = HT; 248 = VT) that (i) harboured variants with a pronounced difference in frequency (±15%) between the two treatments, and (ii) were differentially expressed in the respective treatment. The innermost ring depicts those genes (112 = HT; 8 = VT) from the third ring that were solely differentially expressed in one but not the other treatment. Genome visualization was carried out with Circos (32). (B) Coxcomb chart summarizes the known functional categories of the 112 unique genes differentially expressed in the HT symbiont population. Axis indicates the percentage of genes affected per functional category. All individual values are provided in Dataset S3.

A large contrast in temporal gene regulation

Global gene expression analysis showed a highly dynamic transcriptional landscape during infection, including pronounced temporal signatures consistent with previous findings for chlamydial symbionts (33). Strikingly, the gene expression profiles of VT and HT symbionts were very different at all time points analyzed, indicating marked changes in temporal gene regulation (**Fig. 4*A*** and Dataset S4). To dissect these temporal differences we next focused more specifically on which genes were differentially expressed between two consecutive time points (Dataset S5). We observed the major discrepancy between HT and VT symbionts when comparing the number of differentially expressed genes between released EBs and 2 hpi, reflecting early events during infection (**Fig. 4*B***). In the symbionts from the VT regime, only 137 genes were differentially expressed between these two early time points, whereas 1,025 genes were differentially expressed in the symbionts from the HT regime (**Fig. 4*B***). We next investigated whether the identified differentially expressed genes are associated with the set of genes that were characterized by a major shift in variant frequency in our previous analysis. In fact, from this set of 446 genes, 354 and 248 were differentially expressed at some point during the developmental cycle

of HT and VT symbionts, respectively (**Fig. 3*A*** and Dataset S6). When we narrowed down our search in these two groups to those genes that were solely differentially expressed in one regime, we identified only 8 unique genes for the VT symbionts (Dataset S7). By contrast, a total of 112 unique genes were identified for the HT symbionts (**Fig. 3*A*** and Dataset S7). Given that the latter genes exhibited significant changes in variant frequencies and were solely differentially expressed in HT symbionts, they are likely associated with the increase in virulence in the symbiont population from the HT regime. This notion is further supported by the fact that the function of these genes have been previously implicated in chlamydial survival and virulence (33, 34). This included oxidative phosphorylation, sensing and signal transduction, transport, as well as the type III secretion system (T3SS) (**Fig. 3*B***).

Pronounced differences observed at the single gene and pathway level

How were these and other pathways and functional categories represented in the distinct temporal expression patterns of HT and VT symbionts? To address this, we analyzed gene expression levels and gene regulation of selected pathways in detail. We consistently observed that gene expression levels were broadly similar for both HT and VT symbionts during replication (i.e. at 24 hpi; **Fig. 4*C*** and Dataset S4). The most highly expressed genes at this time point are involved in cellular processes such as cell division and peptidoglycan synthesis, DNA replication and translation, and metabolic activity (lipid and amino acid transport and metabolism), together representing hallmarks of the chlamydial replicative stage (26, 35) (*SI Appendix*, Figs. S1-S3 and Dataset S9). In contrast, pronounced differences between HT and VT chlamydiae at the single gene and pathway level were apparent at the onset of infection (2 hpi) and even more so at the EB stage (**Fig. 4*D***). This included a significant upregulation of structural components, effectors and chaperones of the T3SS early in infection in HT symbionts, while this system was not differentially expressed at this time point in VT symbionts (**Fig. 4*D***). The T3SS was already present in the last common ancestor of all extant chlamydiae and is a primary and evolutionarily well-conserved chlamydial virulence factor (36, 37). Another set of genes upregulated early in HT but not VT functions in fatty acid biosynthesis (**Fig. 4*D***). As obligate intracellular bacteria, chlamydiae rely on subversion of diverse host cell metabolites, yet *de novo* fatty acid and lipid biosynthesis is considered essential

for replication (38). The largest difference in gene expression between HT and VT symbionts was observed in the released EBs, the extracellular stage required for survival and infection of new host cells. HT symbionts showed pronounced upregulation of genes involved in oxidative phosphorylation, glycolysis/gluconeogenesis, the citric acid cycle, pyruvate/propanoate metabolism, and the pentose phosphate pathway - all involved in the utilization of carbon substrates and energy conservation. This pattern was not detectable in EBs from the VT regime (**Fig. 4*D***). However, it resembles the gene expression profile of EBs in closely related chlamydial symbionts for which glucose metabolism and respiration has been shown to be important at the infectious stage (33, 34).

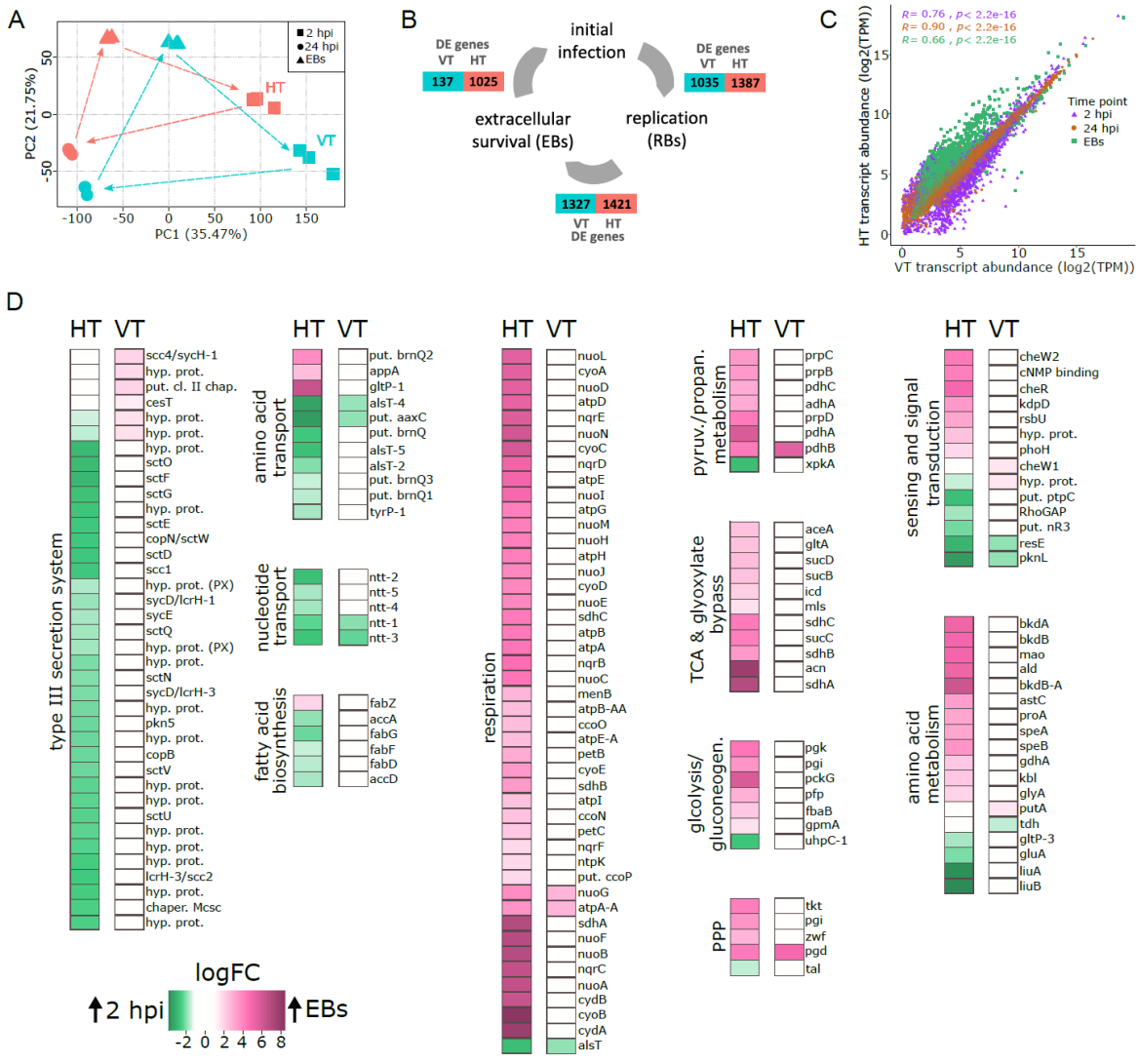

Fig. 4. Genes involved in extracellular survival and virulence were upregulated in HT symbionts. (A) Principal component (PC) analysis plot for RNA-Seq gene expression data at three time points of symbionts from the HT or VT treatment that were collected from amoeba host cells (2 hpi and 24 hpi) or from the medium (released EBs) (*n* = 3). (B) An overview of the total number of differentially expressed (DE) genes for either treatment between two

consecutive time points that marked crucial developmental events during infection (n = 3). Genes were considered differentially expressed if their expression changed twofold with a false-discovery rate (FDR) smaller or equal 0.05. (C) Scatter plot showing the correlation between transcript abundance in the two treatments at each time point (n = 3). Each dot on the scatter plot represents a single gene. Kendall's tau coefficient at each time point is shown on the plot. (D) Heat maps depict those metabolic pathways and T3SS loci with marked differences between the HT and VT treatments with respect to $\log_2$ fold changes (logFC) of genes that were differentially expressed between released EBs and the 2 hpi (false discovery rate of 0.05) (n = 3). A logFC below 1 (no differential expression) is shown in white, significant upregulation in released EBs is shown in pink, and significant upregulation at 2 hpi is shown in green. The stronger the color, the stronger the gene expression change between the two time points. All logFC values are provided in Dataset S8.

Conclusions

Our gene expression analysis supports a scenario in which HT symbionts that invade the host as EBs were better adapted to both survival in the extracellular environment and infection of their amoeba host than VT symbionts. The more virulent HT phenotype as observed in the symbiont fitness assays (**Fig. 2**) can be explained by a more pronounced expression of genes and pathways known to be important for EB maturation and extracellular persistence. Outside the host cell, access to nutrients, such as glucose, and respiration is crucial to maintain infectivity (34, 39, 40) , and it is exactly the genes involved in these processes that are affected by both strong variant shifts and pronounced differential expression (**Figs. 3 and 4**). When encountering a new host cell, secretion of effector proteins is a crucial initial step of infection and required for host cell manipulation to establish the intracellular niche (41). The main mechanism to achieve this, the chlamydial T3SS, is more strongly upregulated in HT symbionts at 2 hpi, providing a likely reason for the significantly higher infectivity in these compared to VT symbionts (**Fig. 2*B***).

VT symbionts greatly resembled the founder ancestral population, with respect to both phenotype and genotype. When shown relative to each other at 6 hpi and at 48 hpi, the infectivity of the VT symbionts did not differ greatly from the infectivity of the ancestral population (**Fig. 2*B* and *C***). Furthermore, the vast majority of ancestral variant frequencies did not change much in the VT population, with novel mutations

largely remaining at low frequencies (**Fig. 3*A*** and Dataset S2). Conversely to the HT symbionts, those genes and pathways known to be essential for extracellular survival and infection were not strongly upregulated in the VT symbionts, and this likely determined their lower infectivity. The attenuated character of the VT population makes the symbionts better suited for long-term coexistence with their amoeba hosts, and this regime more closely resembles the transmission mode of these bacteria under standard laboratory conditions, where the symbionts do not encounter naive hosts.

In nature one expects a mixed transmission mode for these environmental chlamydiae, where both HT and VT co-occur simultaneously. Numerous factors interact together, some of which favor one transmission mode over the other. One such factor likely playing an important role in the environment is host cell density. Where host individuals are common and encountering new hosts is not limiting, HT would increase because selection on host fitness is relaxed. Furthermore, an increase in HT results in higher symbiont diversity within hosts leading to within-host competition which, in turn, promotes more virulence (20, 42). However, at low host cell density VT is more likely to dominate because host fitness is particularly crucial for symbiont survival, selecting for attenuated phenotypes. We argue that in our system, virulence (i.e. host fitness costs due to the infection) is ultimately a compromise between within-host replication and host mortality. In fact, the trade-off hypothesis, discussed and challenged for the past three decades, proposes that transmission will have a strong effect on virulence evolution (43, 44). In the model system investigated here, the symbionts are required to strike an optimal balance between being more benign (VT population) and more parasitic (HT population), and this will vary depending on the complex web of interactions in the surrounding environment.

In conclusion, we have shown that conditions favoring HT led to increased parasitism and higher infectivity in a bacterial symbiont related to major pathogens of humans. While predicted by theory (11, 15–17), experimental evidence for bacteria is sparse (30, 31, 45). Here, the observed shift towards parasitism occurred within a short time frame and was mainly driven by a change in standing genetic variation,

i.e. selection for several hundred SNPs that occurred in the original symbiont population. The parasitic phenotype was characterized by an extensive adjustment of gene expression, mainly affecting functions important at the infectious stage of the symbiont and required for host cell manipulation. The experimental and molecular evidence presented in this study contributes to a better understanding of how conditions affecting transmission mode can lead to the emergence of bacterial parasitism.

Materials and Methods

Evolution experiments.

The same ancestral population of *Acanthamoeba* sp. UWC1 containing *Parachlamydia acanthamoebae* UV-7 was used in the two different regimes. This population had been maintained at 20 °C in TSY medium [30 g/l trypticase soy broth, Oxoid, 10 g/l yeast extract; pH 7.3]. Cultures were regularly screened by fluorescence *in situ* hybridization and 4',6'-diamidino-2-phenylindole (DAPI) staining (0.1 *µ*g/ml) to exclude contamination.

In the VT regime, amoebae that were fully infected with the symbiont were maintained in 25 cm^2 cell culture flasks (Nalge Nunc International, Rochester, NY, USA) with 10 ml TSY medium at 20 °C. The medium was replaced daily, ensuring that as much liquid as possible, containing released EBs and cell debris, was removed. New amoebae were never added to the cultures, and the bacterial symbionts were passed on from parent cell to daughter cell. This was performed for 14 months, equivalent to 525 symbiont generations. Amoeba-chlamydia samples from the end of the VT experiment were stored with dimethylsulfoxide as a cryoprotectant in liquid nitrogen. These samples were later used to isolate symbionts for the fitness assays.

At the start of every week in the HT regime, 3 ml TSY medium, 2 ml of a suspension of naive amoeba cells (the appropriate number of amoebae was pre-determined during the initial infection cycles) and 2 ml of filtered supernatant (containing released EBs) from the previous week's population were mixed together in 25 cm^2 cell culture flasks. The mixture was allowed to stand for 30 min at room temperature, followed by centrifugation (1,000 rpm, 15 min, 23 °C) so as to synchronize the infection. The medium was then exchanged by pipetting, after which the cultures

were allowed to grow for a week at 20 °C. Amoebae were completely infected by the end of the week, with many EBs being released in the medium. Infected amoebae were then discarded and the filtered supernatant was used to infect naive host cells.This was performed for 14 months, equivalent to 560 symbiont generations. EBs from the end of the HT experiment were stored in SPG buffer (75 g/l sucrose, 0.52 g/l KH_2PO_4, 1.53 g/l $NaHPO_4 \cdot 7H_2O$, 1.53 g/l $Na_2HPO_4 \cdot 2H_2O$, 0.75 g/l glutamic acid; pH 7.2) at -80 °C. These EBs were later used in the fitness assays.

Fitness assays.

To measure the effect that the HT and VT symbionts had on growth rate of host cells, amoeba populations that were either fully infected with symbionts isolated from the HT or VT treatment, or uninfected, were seeded in 1 ml TSY medium in 24-well plates (Thermo Fisher Scientific Inc., Waltham, MA, USA) at a concentration of 2 x 10^4 cells per ml and incubated at the HT/VT regime temperature of 20 °C. The growth of the amoebae after 7 days was quantified by counting the number of host cells in a Neubauer chamber (VWR International, Radnor, PA, USA). The initial and final cell numbers were used to calculate the net growth rate per day: $\ln(N_t - N_0)/t$; N_t is the final cell concentration, N_0 is the initial cell concentration, and t is 7 days (46).

To quantify the virulence of the bacteria, naive amoebae were infected with symbionts purified from the HT or VT treatment at the end of the experiment, which was conducted as follows. Culture supernatant was filtered through 1.2 *µ*m syringe filters (Sartorius, Göttingen, Germany) to remove residual host cells. Bacteria were collected by centrifugation (10,000 rpm, 10 min, 4 °C), resuspended in precooled SPG buffer, homogenized using a 21-gauge injection needle (B. Braun, Melsungen, Germany), and stored at -80 °C in SPG buffer. For quantification of purified EBs, 10 *µ*l cell suspensions in 10 ml Page's amoeba saline (PAS) (0.12 g/l NaCl, 0.004 g/l $MgSO_4 \cdot 7H_2O$, 0.004 g/l $CaCl_2 \cdot 2H_2O$, 0.142 g/l Na_2HPO4, 0.136 g/l KH_2PO4) were filtered onto a polycarbonate membrane with a pore size of 0.2 *µ*m (EMD Millipore, Billerica, MA, USA); cells were stained with 0.1 *µ*g/ml DAPI and counted using an epifluorescence microscope (Axioplan 2 imaging; Carl Zeiss, Oberkochen, Germany).

Symbiont-free amoebae were harvested right before infection and an aliquot was counted in a Neubauer chamber. The amoeba cell suspension in TSY was seeded in

the required amount of wells (1 ml/well) in a 24-well plate at a concentration of 10^5 amoebae/ml. The plate was then incubated at 20 °C for 1 hour to allow the host cells to attach prior to infection. EBs for the samples to be tested were thawed at 37 °C in a water bath and added in triplicate to wells in the 24-well plate at a multiplicity of infection (MOI) of 5, as this was predetermined to be the best MOI to observe differences in infectivity between samples. The well-plates were then incubated for 15 min at room temperature, followed by centrifugation (1,000 rpm, 15 min, 23 °C) so as to synchronize the infection. The medium was then changed so as to remove any residual EBs. This was done by removing 1 ml TSY per well by pipetting, and then re-adding fresh 1 ml TSY. Wells were then sampled at 6 hpi and 48 hpi, with the latter time point being divided into supernatant fraction and amoeba fraction. Quantification of symbionts was then carried out by qPCR (see below). The relative infectivity was determined by dividing the total symbiont number per ml of the particular treatment (HT/VT) by that of the ancestral symbiont.

Quantitative PCR (qPCR).

We produced a standard for the absolute quantification of members of the *Chlamydiales* by qPCR assay with the primer pair panCh16F2_mod (5′-CCGCCAACAYTGGGACT-3′) / panCh16R2_mod (5′-GKAGGTRGCCGCYGCTTCTTTAC-3′) (both modified from (47)) targeting the *Chlamydiales* 16S rRNA gene. This qPCR standard consisted of a linear DNA molecule containing the target sequence of the primer pair that was used in the qPCR assays flanked by 90 and 76 base pair regions on each end to allow the qPCR polymerase to attach. Briefly, the standard was produced by amplifying the target sequence by PCR, cloning into the pCR4TOPO vector using the TOPO-TA Cloning Kit for Sequencing by Thermo Fisher Scientific, transforming the resulting plasmids into One Shot TOP10 chemically competent *Escherichia coli* cells and finally purifying and pooling using the QIAquick PCR Purification Kit from Qiagen.

DNA from the amoeba-chlamydia cultures was isolated using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). All qPCR assays were performed in 16 *µ*l, with iQ SYBR Green Supermix (Bio-Rad Laboratories, Inc., Hercules, CA, USA), 0.25 µM concentrations of each primer (both modified from (47)), molecular biology grade water (Merck KGaA, Darmstadt, Germany), and 2 µl of DNA sample using a

CFX96 Touch Real-Time PCR Detection System from Bio-Rad. The cycling conditions were 3 min at 95 °C, followed by 40 cycles of 15 s at 95 °C, 30 s at 60 °C, and 30 s at 72 °C. The results of the qPCR runs were analyzed using the CFX Manager 3.1 from Bio-Rad. The melt curves were checked for any abnormalities. The absolute number of 16S-rRNA genes in the samples was determined by running a serial dilution of the standard sequences of known concentration in every qPCR run and comparing the amplification of the samples to the linear regression (R^2 > 0.99 in all cases) produced by the standards. All samples were run in triplicate and molecular biology grade water was used as a negative PCR control.

DNA extraction and sequencing.

Frozen amoeba-chlamydia cultures from the founder population as well as from the end of the experiment of the VT regime were thawed in a water bath at 35 °C for 2 min and then pipetted into 25 cm^2 culture flasks with 10 ml TSY to allow sufficient growth. After one week, these cultures were then upscaled into 175 cm^2 flasks (Nalge Nunc International, Rochester, NY, USA) for another week so as to increase the quantity of EBs that could be isolated. Frozen EBs from the HT regime at the end of the experiment were used to infect naive amoeba cells and upscaled in the same way as the VT regime. EB purification from the supernatant was done as described previously for both regimes. Isolation of genomic DNA was then carried out on these isolated EBs using a cetyl trimethylammonium bromide-based extraction method (48). Library preparation and pool sequencing of symbiont populations were performed at the Vienna BioCenter Core Facilities (VBCF) Next-Generation Sequencing (NGS) Unit (https://www.viennabiocenter.org/facilities/). Illumina Genome Analyser and HiSeq2000 instruments were used to generate paired-end reads of ~125 bases according to standard procedures, achieving the desirable coverage of 300- to 600-fold.

Sequence read processing and mapping.

The quality filtering and trimming of the Illumina reads was done by Prinseq-lite (v0.19.5) (49) and Trimmomatic (v0.32) (50) and applied as follows. First, a sliding window with size 10 removed any bases with lower quality than 20 starting from the

3’ end by removing the read part containing the low-quality bases. Any reads shorter than 40 nucleotides were also removed. We discarded any low quality read that had an average Phred score below 30. Only read pairs were kept. These reads were mapped against the reference genome sequence of *P. acanthamoebae (27)* (NCBI RefSeq accession number NC_015702.1) using the Burrows-Wheeler Aligner (BWA) (v0.7.5a) (51) with standard settings and stored as bam files. For conversions from sam to bam files and from bam to fastq files (as Cortex_var input), we used SAMtools (v0.1.18) (52) and Picard Tools (v1.92, http://broadinstitute.github.io/picard/).

Variant prediction.

We used our novel VarCap pipeline (https://github.com/ma2o/VarCap) (53) to detect single nucleotide and structural differences at high resolution and accuracy based on Illumina reads mapped to the reference *P. acanthamoebae* genome sequence. VarCap automatically performs quality filtering, mapping, variant calling and post-filtering of the predicted variants. The output is a Variant Call Format (VCF) file with a detailed description of the variants and their relative frequencies as well as two PDF files, which give a graphical overview of variant coverage and their frequency distribution.

Zojer and colleagues separately evaluated numerous variant detection tools for SNPs, insertion and deletions, as well as structural variants. Owing to the different variant calling abilities of the separate tools at low frequencies, several tools were combined so as to increase the sensitivity of the VarCap pipeline. Furthermore, a variant call had to be supported by at least two different tools so as to gain precision and robustness. Combining all these selected software tools, all variants except inversions could be detected at a minimum read abundance of 2% at 400x total coverage, with a minimum of 8 reads per variant, with high sensitivity (53).

RNA-Seq infection experiments.

EBs were freshly purified from amoeba cultures grown in 500 cm^2 culture flasks (Nalge Nunc International, Rochester, NY, USA), in which EBs had been allowed to accumulate in the medium for 1 week. These EBs originated from the end of the

experiment of both the HT and VT regimes. Purification of EBs and subsequent quantification of purified EBs was carried out as described previously. Symbiont-free amoebae were harvested 3 days before infection and seeded at low cell density, followed by incubation at 20 °C until infection. To optimize infection efficiency, particularly at early time points, we used an MOI of 185 for 2 hpi, an MOI of 100 for 24 hpi and an MOI of 25 for released EBs accumulated at 7 days post infection. Precultivated amoebae were harvested, transferred to 50 ml Greiner tubes (Greiner Bio-One GmbH, Frickenhausen, Germany), and purified *Parachlamydia* EBs were added, followed by repeated centrifugation (centrifuged at 130 x g twice for 5 min each time and then once for 10 min at 20 °C) with vortexing between the centrifugation steps. Infected amoebae were then transferred back to the culture flasks and incubated in TSY medium at 20 °C for 2 hr before the infection was synchronised by gently washing the attached amoebae three times with PAS. TSY medium was added to the cultures; some culture flasks were sampled at the 2 hpi time point, whereas the remaining culture flasks were incubated at 20 °C for 24 hr. Released *Parachlamydia* EBs were harvested from the medium supernatant after one week of incubation from amoeba-endosymbiont cultures that were grown at the same conditions, and were purified as described previously. All infection experiments were performed in biological triplicates.

RNA extraction and sequencing.

In order to increase the coverage of the *Parachlamydia* transcriptome, a protocol for enrichment of bacteria prior to RNA extraction was employed (33). Each sample was processed in less than 7 min so as to minimize possible changes of the transcriptomes during enrichment, since the half-life of total mRNA from *E. coli* was demonstrated to be in this range (54). Infected amoebae were harvested and collected (7,600 X g, 2 min, 20 °C), after which the pellets were resuspended in a sucrose buffer (35 mMTris-HCl, 250 mM sucrose, 25 mMKCl, 10 mM $MgCl_2$) supplemented with 50 *μ*g/ml rifampicin (Sigma-Aldrich Handels GmbH, Vienna, Austria) in order to inhibit active transcription during the enrichment procedure (55, 56). Amoebae were then disrupted by vortexing in the presence of glass beads (diameter of 0.75 to 1 mm; Carl Roth) for 1 min. The suspensions were subsequently filtered through a 5 *μ*m filter, the flowthrough fractions containing the bacteria were

collected by centrifugation (10,600 X g, 2 min, room temperature), and the pellets were immediately resuspended in TRIzol reagent (Thermo Fisher Scientific). Extracellular *Parachlamydia* EBs were pelleted (20,800 X g, 2 min, room temperature); these pellets were then resuspended in sucrose buffer and treated like the enriched bacteria.

Cells were mechanically disrupted by bead beating for 30 s at 4.5 m/s using lysing matrix A tubes and a FastPrep-24 instrument (MP Biomedicals, Santa Ana, CA, USA). Subsequent RNA extraction was performed according to the TRIzol guidelines. Residual DNA was then digested using the Turbo DNA-free kit (Thermo Fisher Scientific) as recommended by the manufacturer. DNase-treated RNA was precipitated with sodium acetate and ethanol and dissolved in nuclease-free water (Thermo Fisher Scientific), whereas DNA contamination was controlled for via PCR targeting a short region of the bacterial 16S rRNA gene (SigF2/R2 primers (57)) using 35 PCR cycles. rRNA was subsequently removed using the Ribo-Zero magnetic kit for Gram-positive bacteria according to the manufacturer's instructions (Illumina, San Diego, CA, USA). For strand-specific cDNA library preparation, the NEBNext Ultra directional RNA library prep kit for Illumina, in combination with the NEBNext multiplex oligonucleotides (New England Biolabs, Ipswich, MA, USA), was used starting at first-strand cDNA synthesis. All libraries were sequenced using an Illumina HiSeq2500 system at the VBCF NGS Unit with 50 bp read length.

RNA-Seq read processing.

Sequencing reads were trimmed and cleaned before mapping (Dataset S10), based on sequencing read statistics obtained using Trimmomatic (v0.32) (50) and FastQC (v0.10.0) (58). Briefly, (i) the first 11 bases were removed, (ii) any remaining adapter sequences were clipped, and (iii) sequence reads shorter than 25 bases were removed, all done using Trimmomatic. To map bacterial reads to the *Parachlamydia acanthamoebae* reference genome (27) BWA (v0.7.5a) (51) was used. Only unambiguously mapped reads were kept using SAMtools (v0.1.18) (52). Strand-specific reads per predicted gene were counted via HTSeq (v0.11.1) (59).

Gene expression analyses.

Differentially expressed genes were determined between two consecutive time points (2 hpi to 24 hpi, 24 hpi to EBs, EBs to 2 hpi) using the R software environment (v3.6.1) and the Bioconductor package DESeq2 (v1.22.1) (60–62). Genes were considered differentially expressed if their expression changed twofold with a false-discovery rate (FDR) smaller or equal 0.05. Log_2 fold changes (logFC) of differentially expressed genes and mean centered gene expression values (log_2 transcripts per kilobase million [TPM]) were used for visualization as heat maps using the R package gplots (v3.0.1.1) (63).

To improve the available *Parachlamydia acanthamoebae* genome annotation, for each gene we collected Pfam domains (bit score $\geq$ 25) (64), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway maps (65), gene ontology (GO) terms using Blast2GO (66), clusters of orthologous groups of proteins (COGs) and their functional categories using MaGe (67), and type III secretion effector predictions using Effective (68). The Bioconductor software package GOseq (69) was used to test for statistical enrichment of functional categories among different gene sets.

Statistics.

Statistical analysis was performed using R (v3.6.1) (60). Details on sample sizes, in addition to the statistical tests conducted, are shown in the corresponding figure legends; the statistical parameters and significance are reported in the figure legends. Data were considered to be statistically significant when $P \leq 0.05$. The fitness assays were performed on three to eight replicates, consisting of three technical replicates each for qPCR data. Comparisons for two groups were calculated by unpaired two-tailed Student's *t* tests and the comparison for more than two groups was calculated by a Kruskal-Wallis test followed by Dunn's multiple comparison post-hoc test.

Three infection experiments (representing three time points) using different symbiont sub-populations were performed, and three samples for each time point and treatment (VT and HT) were used for RNA sequencing. We performed a number of statistical tests on the RNA-Seq data obtained in triplicate: a negative binomial

model followed by a Wald test in DESeq2 (62) was used to identify differentially expressed genes between two consecutive time points for each treatment; a Probability Weighting Function followed by a Wallenius approximation in GOseq (69) was used to test for statistical enrichment of functional categories (FDR $\leq$ 0.05) among different gene sets for each treatment; two-tailed Fisher's exact tests were conducted to test whether predicted type III secreted proteins were significantly enriched in any given gene set for each treatment; and Kendall rank correlation was used to test the correlation between transcript abundance in the two treatments.

Data availability.

All raw sequencing data was deposited in the National Center for Biotechnology Information (NCBI) database under bioproject accession number PRJNA574613. RNA-Seq data can be accessed from the Gene Expression Omnibus (GEO) database through accession number GSE138099.

Acknowledgments

We thank Christian Kost for providing very helpful comments on the manuscript. Sequencing was performed at the Vienna Biocenter Core Facility (VBCF) NGS Unit (https://www.viennabiocenter.org/facilities/).

This project has received funding from the European Research Council ERC (EVOCHLAMY, grant no. 281633 to MH) and the Austrian Science Fund FWF in the context of the doc.funds program (DOC 69-B) and the InfectERA project Eugenpath (project number I1628-B22).

MH conceived the study. PH, LS, LK, CW and MH designed the experiments. LS, JS, FW and HN performed the evolution experiment. PH and TK performed the fitness assays. PH, CW, and JS performed the genome and transcriptome sequencing experiments. PH, LK, MZ and MH analysed fitness, genome and transcriptome data. SK and TR provided data analysis support. PH and MH wrote the manuscript. PH, LS, CW, LK, SK, MZ, TR, and MH edited the manuscript.

Correspondence and material requests should be addressed to MH (matthias.horn@univie.ac.at).

References

1. P. W. Ewald, Transmission modes and evolution of the parasitism-mutualism continuum. *Ann. N. Y. Acad. Sci.* **503**, 295–306 (1987).

2. Y. Michalakis, I. Olivieri, F. Renaud, M. Raymond, Pleiotropic action of parasites: How to be good for the host. *Trends Ecol. Evol.* **7**, 59–62 (1992).

3. J. D. Hoeksema, *et al.*, A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. *Ecol. Lett.* **13**, 394–407 (2010).

4. J. L. Sachs, R. G. Skophammer, J. U. Regus, Evolutionary transitions in bacterial symbiosis. *Proc. Nat. Acad. Sci.* **108**, 10800–10807 (2011).

5. M. Lipsitch, S. Siller, M. A. Nowak, The evolution of virulence in pathogens with vertical and horizontal transmission. *Evolution* **50**, 1729–1741 (1996).

6. N. Yamamura, Evolution of mutualistic symbiosis: A differential equation model. *Res. Popul. Ecol.* **38**, 211–218 (1996).

7. M. Bright, S. Bulgheresi, A complex journey: transmission of microbial symbionts. *Nat. Rev. Microbiol.* **8**, 218–230 (2010).

8. S. A. Frank, Models of symbiosis. *Am. Nat.* **150** (suppl. 1), S80–99 (1997).

9. N. Yamamura, Vertical Transmission and Evolution of Mutualism from Parasitism. *Theor. Popul. Biol.* **44**, 95–109 (1993).

10. N. A. Moran, J. P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* **42**, 165–190 (2008).

11. R. Axelrod, W. D. Hamilton, The evolution of cooperation. *Science* **211**, 1390–1396 (1981).

12. R. M. Anderson, R. M. May, Coevolution of hosts and parasites. *Parasitology* **85**, 411–426 (1982).

13. J. J. Bull, W. R. Rice, Distinguishing mechanisms for the evolution of co-operation. *J. Theor. Biol.* **149**, 63–74 (1991).

14. P. E. Turner, V. S. Cooper, R. E. Lenski, Tradeoff between horizontal and vertical modes of transmission in bacterial plasmids. *Evolution* **52**, 315–329 (1998).

15. P. E. M. Fine, Vectors and vertical transmission: an epidemiologic perspective. *Ann. NY Acad. Sci.* **266**, 173–194 (1975).

16. P. W. Ewald, Host-Parasite Relations, Vectors, and the Evolution of Disease Severity. *Annu. Rev. Ecol. Syst.* **14**, 465–485 (1983).

17. J. J. Bull, Perspective: Virulence. *Evolution* **48**, 1423 (1994).

18. R. M. May, R. M. Anderson, Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. Lond. B Biol. Sci.* **219**, 281–313 (1983).

19. J. J. Bull, I. J. Molineux, W. R. Rice, Selection of Benevolence in a Host-Parasite

System. *Evolution* **45**, 875 (1991).

20. E. A. Herre, Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science* **259**, 1442–1445 (1993).

21. P. W. Ewald, *Evolution of Infectious Disease*, (Oxford University Press, 1994).

22. S. L. Messenger, I. J. Molineux, J. J. Bull, Virulence evolution in a virus obeys a trade off. *Proc. R. Soc. Lond. B Biol.Sci.* **266**, 397–404 (1999).

23. A. D. Stewart, J. M. Logsdon Jr, S. E. Kelley, An empirical study of the evolution of virulence under both horizontal and vertical transmission. *Evolution* **59**, 730–739 (2005).

24. J. L. Sachs, T. P. Wilcox, A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc. Biol. Sci.* **273**, 425–429 (2006).

25. M. Horn, *et al.*, Illuminating the evolutionary history of Chlamydiae. *Science* **304**, 728–730 (2004).

26. Y. M. Abdelrahman, R. J. Belland, The chlamydial developmental cycle. *FEMS Microbiol. Rev.* **29**, 949–959 (2005).

27. A. Collingro, *et al.*, Unity in Variety--The Pan-Genome of the *Chlamydiae*. *Mol. Biol. Evol.* **28**, 3253–3270 (2011).

28. M. Horn, *Chlamydiae* as symbionts in eukaryotes. *Annu. Rev. Microbiol.* **62**, 113–131 (2008).

29. D. Ebert, The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. *Annu. Rev. Ecol. Evol. S.* **44**, 623–643 (2013).

30. H. Magalon, T. Nidelet, G. Martin, O. Kaltz, Host growth conditions influence experimental evolution of life history and virulence of a parasite with vertical and horizontal transmission. *Evolution* **64**, 2126–2138 (2010).

31. E. Dusi, C. Gougat-Barbera, T. U. Berendonk, O. Kaltz, Long-term selection experiment produces breakdown of horizontal transmissibility in parasite with mixed transmission mode. *Evolution* **69**, 1069–1076 (2015).

32. M. Krzywinski, *et al.*, Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

33. L. König, *et al.*, Biphasic Metabolism and Host Interaction of a Chlamydial Symbiont. *mSystems* **2** (2017).

34. A. Omsland, J. Sager, V. Nair, D. E. Sturdevant, T. Hackstadt, Developmental stage-specific metabolic and transcriptional activity of *Chlamydia trachomatis* in an axenic medium. *P. Natl. Acad. Sci.* **109**, 19781–19785 (2012).

35. A. P. Mäurer, A. Mehlitz, H. J. Mollenkopf, T. F. Meyer, Gene expression profiles of *Chlamydophila pneumoniae* during the developmental cycle and iron depletion-mediated persistence. *PLoS Pathog.* **3**, e83 (2007).

36. H. J. Betts-Hampikian, K. A. Fields, The Chlamydial Type III Secretion

Mechanism: Revealing Cracks in a Tough Nut. *Front. Microbiol.* **1** (2010).

37. A. Subtil, A. Collingro, M. Horn, Tracing the primordial *Chlamydiae*: extinct parasites of plants? *Trends Plant Sci.* **19**, 36–43 (2014).

38. J. Yao, *et al.*, Type II fatty acid synthesis is essential for the replication of *Chlamydia trachomatis*. *J. Biol. Chem.* **289**, 22365–22376 (2014).

39. B. S. Sixt, *et al.*, Metabolic Features of *Protochlamydia amoebophila* Elementary Bodies – A Link between Activity and Infectivity in *Chlamydiae*. *PLoS Pathog.* **9**, e1003553 (2013).

40. A. Omsland, B. S. Sixt, M. Horn, T. Hackstadt, Chlamydial metabolism revisited: interspecies metabolic variability and developmental stage-specific physiologic activities. *FEMS Microbiol. Rev.* **38**, 779–801 (2014).

41. C. Elwell, K. Mirrashidi, J. Engel, *Chlamydia* cell biology and pathogenesis. *Nat. Rev. Microbiol.* **14**, 385–400 (2016).

42. D. Ebert, K. L. Mangin, The influence of host demography on the evolution of virulence of a microsporidian gut parasite. *Evolution* **51**, 1828–1837 (1997).

43. S. Alizon, A. Hurford, N. Mideo, M. Van Baalen, Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evolution. Biol.* **22**, 245–259 (2009).

44. M. A. Acevedo, F. P. Dillemuth, A. J. Flick, M. J. Faldyn, B. D. Elderd, Virulence-driven trade-offs in disease transmission: A meta-analysis. *Evolution* **73**, 636–647 (2019).

45. K. W. Jeon, T. I. Ahn, Temperature sensitivity: a cell character determined by obligate endosymbionts in amoebas. *Science* **202**, 635–637 (1978).

46. A. K. Minks, P. Harrewijn, *Aphids: Their Biology, Natural Enemies, and Control*, (Elsevier Science Limited, 1987).

47. J. Lienard, *et al.*, Development of a new *Chlamydiales*-specific real-time PCR and its application to respiratory clinical samples. *J. Clin. Microbiol.* **49**, 2637–2642 (2011).

48. S. Schmitz-Esser, *et al.*, The genome of the amoeba symbiont "*Candidatus* Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* **192**, 1045–1057 (2010).

49. R. Schmieder, R. Edwards, Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).

50. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

51. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

52. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*

25, 2078–2079 (2009).

53. M. Zojer, *et al.*, Variant profiling of evolving prokaryotic populations. *PeerJ* **5**, e2997 (2017).

54. D. W. Selinger, R. M. Saxena, K. J. Cheung, G. M. Church, C. Rosenow, Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* **13**, 216–223 (2003).

55. U. Dreses-Werringloer, I. Padubrin, H. Zeidler, L. Köhler, Effects of azithromycin and rifampin on *Chlamydia trachomatis* infection in vitro. *Antimicrob. Agents Chemother.* **45**, 3001–3008 (2001).

56. I. Sarov, Y. Becker, Deoxyribonucleic acid-dependent ribonucleic acid polymerase activity in purified trachoma elementary bodies: effect of sodium chloride on ribonucleic acid transcription. *J. Bacteriol.* **107**, 593–598 (1971).

57. S. Haider, A. Collingro, J. Walochnik, M. Wagner, M. Horn, *Chlamydia*-like bacteria in respiratory samples of community-acquired pneumonia patients. *FEMS Microbiol. Lett.* **281**, 198–202 (2008).

58. S. Andrews, FastQC: A quality control tool for high throughput sequence data. FastQC (2010). http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 22 November 2018.

59. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

60. R Development Core Team., R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2011).

61. R. C. Gentleman, *et al.*, Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5** (2004).

62. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

63. G. Warnes, *et al.*, gplots: various R programming tools for plotting data. R package version 2. R Foundation for Statistical Computing, Vienna, Austria (2005).

64. S. El-Gebali, *et al.*, The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).

65. M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, K. Morishima, KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).

66. S. Götz, *et al.*, High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **36**, 3420–3435 (2008).

67. D. Vallenet, MaGe: a microbial genome annotation system supported by synteny results. *Nucleic Acids Res.* **34**, 53–65 (2006).

68. M.-A. Jehl, R. Arnold, T. Rattei, Effective—a database of predicted secreted

bacterial proteins. *Nucleic Acids Res.* **39**, D591–5 (2011).

69. M. D. Young, M. J. Wakefield, G. K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).

Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont

Paul Herrera, Lisa Schuster, Cecilia Wentrup, Lena König, Thomas Kempinger, Hyunsoo Na, Jasmin Schwarz, Stephan Köstlbacher, Florian Wascher, Markus Zojer, Thomas Rattei, Matthias Horn

Supplementary information

SI Text

Calculating the number of symbiont generations in each regime.

In order to put the observed evolutionary changes into context, it was important to calculate the total number of symbiont generations over the course of the experiment as best as possible. This would allow us to compare the number of generations between HT and VT regimes, since a large difference between the two values may improve our understanding of any observed changes.

In the VT treatment, TSY medium was replaced daily and this removed any released EBs present in the medium. Although these EBs were removed from the flask, thus preventing their ability to infect host cells, they still contributed to the total number of symbiont generations. We quantified the amount of EBs present in the medium per day over the course of three successive days with qPCR. We also quantified the amount of symbionts present in a fully infected continuous culture with freshly replaced medium with qPCR. In this way we were able to calculate the number of generations arising from released EBs per day. Apart from this, we measured the growth rate of amoeba infected with symbionts from this VT treatment. Since bacteria are transmitted vertically from parent to daughter cell, the doubling of amoeba cells would be equivalent to doubling of bacteria. We then added the number of generations arising from the released EBs to the number of generations arising from host cell division to calculate the total number of symbiont generations for the VT regime, which equated to 525 generations for the entire duration of the evolution experiment.

In the HT treatment, symbionts isolated from the medium were added weekly to naive amoebae. By quantifying the number of symbionts added at the start of the week and the number of symbionts in the medium at the end of the week, both with qPCR, we were able to calculate the number of symbiont generations arising from released EBs throughout the week. We also measured the growth rate of amoeba infected with symbionts from this HT treatment. By adding the two values together we were able to calculate the total number of symbiont generations for the HT regime, which equated to 560 generations for the entire duration of the evolution experiment.

Thus, we can say that the observed evolutionary changes in the HT and VT regimes have not simply been brought about as a result of a large difference between the total numbers of symbiont generations.

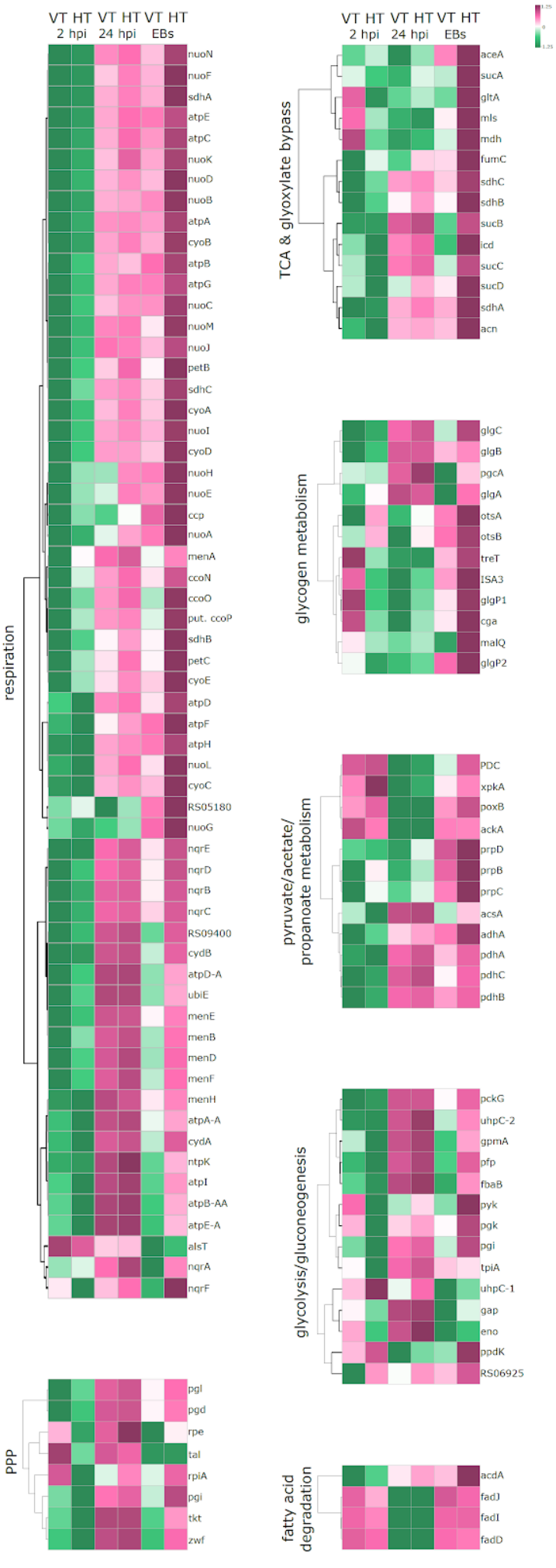
VT HT VT HT VT HT
2 hpi 24 hpi EBs
respiration
nuoN
nuoF
sdhA
atpE
atpC
nuoK
nuoD
nuoB
atpA
cyoB
atpB
atpG
nuoC
nuoM
nuoJ
petB
sdhC
cyoA
nuoI
cyoD
nuoH
nuoE
ccp
nuoA
menA
ccoN
ccoO
put. ccoP
sdhB
petC
cyoE
atpD
atpF
atpH
nuoL
cyoC
RS05180
nuoG
nqrE
nqrD
nqrB
nqrC
RS09400
cydB
atpD-A
ubiE
menE
menB
menD
menF
menH
atpA-A
cydA
ntpK
atpI
atpB-AA
atpE-A
alsT
nqrA
nqrF
PPP
pgl
pgd
rpe
tal
rpiA
pgi
tkt
zwf
VT HT VT HT VT HT
2 hpi 24 hpi EBs
TCA & glyoxylate bypass
aceA
sucA
gltA
mls
mdh
fumC
sdhC
sdhB
sucB
icd
sucC
sucD
sdhA
acn
glycogen metabolism
glgC
glgB
pgcA
glgA
otsA
otsB
treT
ISA3
glgP1
cga
malQ
glgP2
pyruvate/acetate/
propanoate metabolism
PDC
xpkA
poxB
ackA
prpD
prpB
prpC
acsA
adhA
pdhA
pdhC
pdhB
glycolysis/gluconeogenesis
pckG
uhpC-2
gpmA
pfp
fbaB
pyk
pgk
pgi
tpiA
uhpC-1
gap
eno
ppdK
RS06925
fatty acid
degradation
acdA
fadJ
fadI
fadD

Fig. S1. Heat maps of pathways involved in central carbon metabolism and energy generation.

Heat maps of RNA-Seq data from mean centered gene expression (TPM) of symbiont genes (listed in Dataset S4) in HT and VT regimes at all three time points during infection. All known loci involved or putatively involved in the pathways shown are depicted. Gene names or locus tags (without the prefix PUV_) are indicated. Highly expressed genes are shown in pink, whereas low gene expression is shown in green. The stronger the color, the stronger the gene expression value.


VT HT VT HT VT HT
2 hpi 24 hpi EBs
amino acid transport and metabolism
put. brnQ
alsT-4
alsT-5
aspC
iscS
alsT-2
put. brnQ1
put. aaxC
glnA
put. brnQ3
tyrP-1
gltP-3
alsT-3
tdh
alsT-1
putP
dppD
dppF
glyA
tyrP-4
tyrP-3
gcvH
proC
dppB
gcvPA
proB
gluA
RS04255
glnQ
tyrP-2
kbl
ldh
gltP-2
asnB-2
RS12690
RS06995
tyrP-5
gcvT
RS12685
ald
asnB1
appA
speB
RS09865
RS10575
gcvPB
bkdB
putA
mao
bkdA
bkdB-A
gltP-1
proA
liuB
liuA
astD
RS08285
speA
gdhA
astC
put. brnQ2
liuC
hutH
astB
hutG
RS05190
hutU
hutI
metC
astA
cbs
VT HT VT HT VT HT
2 hpi 24 hpi EBs
cell division and peptidoglycan synthesis
mreC
glmM
tolQ
ftsX
uppS
PBP6
tolB
dapA
glmS
dapB
glmU
murI
murF
ask
lytC
mraY
murE
dapL
ftsH
pal
murA
dapF
lytF
PBP3
murG
murD
mreB
cpg2
asd
murB
murC/ddl
ftsK
tolR2
murF/alr
rodA
ftsW
uppP
PBP2
tolA
rodZ
tolR1
ftsY
put. lysM
murJ
fatty acid biosynthesis
accA
accD
fabF
fabG
fabD
fabH
accC1
fabZ
fabI
accB
nulceotide transport
ntt-5
ntt-1
ntt-3
ntt-4
ntt-2

Fig. S2. Heat maps of metabolic pathways strongly upregulated at the peak of RB activity (i.e. at 24 hpi).

Heat maps of RNA-Seq data from mean centered gene expression (TPM) of symbiont genes (listed in Dataset S4) in HT and VT regimes at all three time points during infection. All known loci involved or putatively involved in the pathways shown are depicted. Gene names or locus tags (without the prefix PUV_) are indicated. Highly expressed genes are shown in pink, whereas low gene expression is shown in green. The stronger the color, the stronger the gene expression value.


VT HT VT HT VT HT
2 hpi 24 hpi EBs
type III secretion system
RS04590
sctL
sctJ-1
sctC-2
sctC-1
RS02625
sctR
sctS
sctT
RS03955
RS00690
RS02635
sycD/lcrH-2
RS03940
RS03950
RS02320
lcrH-3/scc2
RS02325
sycD/lcrH-3
sycD/lcrH-1
RS02330
copB
RS02350
RS02355
sctN
RS02395
RS02455
sctO
scc1
sctV
sctQ
RS09410
copN/sctW
RS02590
RS02605
sctU
pkn5
sctD
sctF
sctE
sctG
RS02435
RS02440
RS02450
RS02415
scc4/sycH-2
RS02405
RS02410
cesT
scc4/sycH-1
RS09170
RS02425
RS02400
sycE
lcrH-1/scc3
RS02420
sctJ-2
VT HT VT HT VT HT
2 hpi 24 hpi EBs
sensing and signal transduction
RS05455
RS10640
RS08085
rsbU
RS05190
RS04700
RS08095
RS09185
RS08080
RS08120
kdpD
cheR
RS07285
phoH
cheW2
RS08090
resE
put. ptpC
pknL
put. nR3
RS00545
RS02900
atoS
cheW1
atoC
chxR
cheB
cheY
RS01955
RS05460
btrV
RS09910
RS09180
cheA
RS13085

Fig. S3. Heat maps of pathways involved in infection, sensing and signal transduction pathways.

Heat maps of RNA-Seq data from mean centered gene expression (TPM) of symbiont genes (listed in Dataset S4) in HT and VT regimes at all three time points during infection. All known loci involved or putatively involved in the pathways shown are depicted. Gene names or locus tags (without the prefix PUV_) are indicated. Highly expressed genes are shown in pink, whereas low gene expression is shown in green. The stronger the color, the stronger the gene expression value.

Chapter IV

Genome dynamics and temperature adaptation during experimental evolution of obligate intracellular bacteria

Title:

Genome dynamics and temperature adaptation during experimental evolution of obligate intracellular bacteria

Authors:

Paul Herrera[1], Lisa Schuster[1], Markus Zojer[1], Hyunsoo Na[1,2], Jasmin Schwarz[1], Florian Wascher[1], Thomas Kempinger[1], Andreas Regner[1], Thomas Rattei[1], Matthias Horn[1]

Author affiliations:

[1] Centre for Microbiology and Environmental Systems Science, University of Vienna

[2] Current address: Department of Energy Joint Genome Institute, Walnut Creek, CA, USA

Corresponding author:

Univ.-Prof. Dr. Matthias Horn

University of Vienna

Department of Microbiology and Ecosystem Science

Althanstr. 14

A-1090 Wien

Austria

phone: +43 1 4277 91208

email: matthias.horn@univie.ac.at

Abstract

Evolution experiments with free-living microbes have radically improved our understanding of genome evolution and how microorganisms adapt to different environments. Yet there is a paucity of such research focusing on strictly host-associated bacteria, even though they are so widespread in nature. Here, we used the *Acanthamoeba* symbiont *Protochlamydia amoebophila*, a distant relative of the human pathogen *Chlamydia trachomatis* and representative of a large group of ubiquitous protist-associated environmental chlamydiae, as a model to study how obligate intracellular symbionts made the pivotal evolutionary leap from protists to multicellular eukaryotes with respect to adaptation to elevated temperature. We established replicate populations under two temperatures (20 °C, 30 °C) and two dilutions (0.1%, 10%) for 420 to 510 generations (38 months). We then used a combination of infectivity assays and pooled whole-genome re-sequencing to identify any evolved phenotypes and the molecular basis of adaptation in these bacteria. Despite an overall reduction in infectivity of the symbionts that evolved at 30 °C, symbionts were better adapted to infect at the temperature they were maintained in, be it ambient or elevated. We identified evidence for positive selection at the genomic level for both temperatures, owing to the increase in the proportion of nonsynonymous mutations relative to synonymous mutations. This was most pronounced at the elevated temperature, where numerous variants persisted throughout multiple time points and reached high frequencies. In addition, small indels (< 4 nt) were overrepresented at 30 °C. Between the two temperature regimes the genes affected by mutations strongly differed with respect to their functional roles. Our results suggest that temperature adaptation was facilitated through attenuation of symbiont infectivity as a trade-off to reduce host cell burden.

Introduction

Nature is replete with bacteria that reside within eukaryotic cells. Examples of intracellular bacteria are present throughout the tree of life, and they are found in virtually all eukaryotes, from protists to fungi, plants and animals. Of special mention are the Candidate Phyla Radiation (Brown et al. 2015), diverse groups of bacteria, which are largely defined based on metagenome assembled genomes but are predicted to be host-dependent due to their restricted metabolic capabilities (Gong et

al. 2014; Castelle and Banfield 2018). Intracellular bacteria reside directly in the cytoplasm of the host or else in host-derived vacuoles (Ray et al. 2009), although exceptions do exist (reviewed in (Schulz and Horn 2015). Host-associated microbes shape the biology, ecology and evolution of their hosts in profound ways, affecting processes and functions including development (Braendle et al. 2003) and nutrition (Bäckhed et al. 2005), as well as immunity (Macdonald and Monteleone 2005) and speciation (Hurst and Werren 2001).

The intracellular lifestyle of bacteria is ancient (Moran et al. 2008; Subtil et al. 2014). The interplay of microbes with predatory protists provided a training ground for the emergence of mechanisms to manipulate eukaryotic hosts, to survive phagocytosis, and to eventually exploit eukaryotic cells as an ecological niche (Molmeret et al. 2005). The documented ability of many bacterial pathogens of humans to thrive in amoebae and the conservation of mechanisms for interference with phagocytosis in protists and macrophages are evidence of these primordial bacteria-host relationships (Strassmann and Shu 2017). Still, the transition from protists to multicellular eukaryotic hosts is a pivotal evolutionary step, which remains largely unexplored.

One group of bacteria that has made this transition early on during evolution and is now associated with a broad range of eukaryotic hosts is the phylum *Chlamydiae* (Collingro et al. 2020). The chlamydiae are well-known for being successful bacterial pathogens of humans, with two members in particular. *Chlamydia pneumoniae* is responsible for as many as 10% of community-acquired pneumonia cases (Grayston et al. 1993; Burillo and Bouza 2010) and has been associated with chronic diseases such as asthma, atherosclerosis as well as Alzheimer's disease (Balin et al. 1998; Campbell and Kuo 2004; Shima et al. 2010). *C. trachomatis* comprises more than 15 serologically defined variants and causes over 100 million infections worldwide every year (Wright et al. 2008; Roulis et al. 2013; Dal Conte et al. 2014). Apart from humans, chlamydiae also infect a large variety of animals, including fish, birds and mammals, causing numerous clinically important diseases in economically significant livestock species such as cattle, goats, sheep, and pigs (DeGraves et al. 2003; Polkinghorne et al. 2009; Schautteet and Vanrompay 2011). But not all chlamydiae are so sinister in nature. The so-called environmental chlamydiae (Horn 2008; Collingro et al. 2020) are found ubiquitously in the environment and the majority of them are considered to be symbionts of

free-living amoebae or other protists, with a limited ability to multiply in non-protozoan hosts (Maurin et al. 2002; Greub et al. 2003; Collingro, Poppert, et al. 2005; Casson et al. 2006; Kebbi-Beghdadi et al. 2011; Sixt et al. 2012).

Chlamydiae most likely represent one of the most ancient groups of strictly intracellular bacteria. Comparative and phylogenetic genome sequence analysis suggested that the last common ancestor of all extant chlamydiae already resided within a eukaryotic host, very likely an ancient protist, several hundreds of millions of years ago (Horn et al. 2004; Subtil et al. 2014; Taylor-Brown et al. 2015; Taylor-Brown et al. 2020). They all share a characteristic biphasic developmental cycle comprising two distinct morphological and physiological stages: the elementary body (EB) survives the extracellular environment and infects new host cells, whereas the reticulate body (RB) replicates inside a host-derived vacuole (Abdelrahman and Belland 2005; Horn 2008). Chlamydiae possess small, reduced genomes and lack essential biosynthetic pathways and therefore rely on numerous metabolites from their eukaryotic hosts (Omsland et al. 2014). Chlamydiae also interact with diverse host cellular pathways and organisms (Elwell et al. 2016; Fischer and Rudel 2018). Owing to the intimate relationships with their host cells, like many strictly intracellular microbes, chlamydiae have so far eluded cultivation in host-free media and are thus challenging to study. Here we used a combination of evolution experiments with phenotype assays and genome re-sequencing to investigate host adaptation of chlamydiae.

Experimental evolution has been of utmost importance to the field of evolutionary biology, from testing evolutionary theories to studying trade-offs and constraints to environmental adaptation (Kawecki et al. 2012). Experimental evolution refers to the study of evolutionary processes in response to imposed controlled conditions in real time. Over the last decades, this approach has become particularly popular using microbes due to a multitude of benefits that make them ideal subjects (Elena and Lenski 2003; Hegreness and Kishony 2007; Buckling et al. 2009; Brockhurst et al. 2010; Wichman and Brown 2010; Dettman et al. 2012; Rainey et al. 2017; Cvijović et al. 2018; Van den Bergh et al. 2018). This includes short generation times, large population sizes achieved in very small volumes, as well as comparatively easy propagation of replicate populations. Among the plethora of evolution experiments, one of the most famous is the long-term evolution experiment (LTEE) initiated by Richard Lenski in 1988 using 12 serially propagated

Escherichia coli populations in minimal medium (Lenski et al. 1991), that is still running to this day. Over eighty studies on this long-term experiment alone have yielded invaluable insights into evolutionary dynamics, fitness trajectories and other details of evolutionary processes (Barrick et al. 2009; Wiser et al. 2013; Lenski et al. 2015; Maddamsetti et al. 2015; Tenaillon et al. 2016; Good et al. 2017). Yet, despite the usefulness of this approach and the multitude of studies on free-living microbes, strictly intracellular bacteria have rarely been investigated using evolution experiments due to the challenges inherent to host-dependent systems. The chlamydiae therefore provide us with a distinct opportunity to gain an initial understanding of the evolutionary dynamics in obligate intracellular bacteria (Joseph et al. 2012; Borges et al. 2013; Herrera et al. 2020). In addition to this, evolution experiments should be able to shed light on the process of temperature adaptation in these bacteria, which was a crucial step in the transition from protist symbionts to chlamydial pathogens of humans and animals.

Here, we have established replicate populations of the environmental chlamydia *Protochlamydia amoebophila* thriving in *Acanthamoeba* hosts under two temperatures (20 °C, 30 °C) and two dilutions (0.1%, 10%) for 420 to 510 generations (38 months). Infectivity assays were set up to identify any evolved phenotypes among the long-term symbiont cultures, and we performed pooled whole genome re-sequencing at selected time points coupled with variant detection. We observed a temperature-dependent decreased infectivity of symbionts in the evolved populations and, by tracking genome dynamics, evidence for positive selection. Overall, we provide data with respect to the phenotypic response and the molecular basis of adaptation to the elevated temperature.

Materials and Methods

Evolution experiment

The evolution experiment was founded when ancestral *A. castellanii* Neff was freshly infected with *P. amoebophila* UWE25 (Collingro, Toenshoff, et al. 2005). After a continuous culture was established, the experiment was set up in 24-well plates (Thermo Fisher Scientific Inc., Waltham, MA, USA) under two temperatures (20 °C and 30 °C) and two dilutions (0.1% and 10%). Each treatment had 12 replicates with 1 ml TSY (30 g/l trypticase soy broth, Oxoid, 10 g/l yeast extract; pH 7.3) in each well. Populations were propagated for 38 months by weekly 10-fold dilutions in 1 ml

TSY in the 10% dilution treatment (allowing for ~3.35 generations per week) and monthly 1,000-fold dilutions in 1 ml TSY in the 0.1% dilution treatment (allowing for ~11 generations per month). Samples were taken every 2–3 months and stored with DMSO as a cryoprotectant in liquid nitrogen, where they were available for genome and phenotype analyses later on.

Calculation of generation times

Unlike in free-living bacteria where the generation times can easily be established, this is not as straightforward for obligate intracellular bacteria since they are confined within their host. We needed to account for two aspects to estimate bacterial generation times – (i) *vertical transmission* in the host; and (ii) *horizontal transmission* through release of EBs from the host and infection of a new host. The total number of bacterial generations then equals the sum of the generations in both transmission modes.

To determine this, we first calculated amoeba host generation numbers. In the 10% dilution treatment, where we performed weekly transfers, the number of generations of amoebae per week was calculated by counting aliquots of amoebae in a Neubauer chamber from a particular well in the 24-well plate at the start and at the end of the week. The number of generations was then calculated using the formula $n = \log (N) – \log (N_0) / 0.301$, where n is the number of generations, N_0 is the initial number of amoebae and N is the final number of amoebae after a week. The same was done for the 0.1% dilution treatment, but the counts were carried out after a month as in this case the transfers were monthly.

Temperature had no significant effect on the amoeba generation times at either dilution, most likely due to the finite number of amoebae a well can contain. The mean number of generations per week in the 10% dilution treatment was calculated to be 3.2 and the mean number of generations per month in the 0.1% dilution treatment was 10. These values equate to the vertical transmission of the bacteria. Owing to the fact that the amoebae are fully infected in a continuous culture, when the host cell divides, the bacteria inside will roughly divide between the two daughter cells, with each bacterium needing to divide once so as to reach the maximum bacterial number possible in a fully infected amoeba. Thus, amoeba generation number is equal to bacteria generation number if one considers solely vertical transmission.

So as to assess the extent of horizontal transmission of the bacteria, it was also required to measure the amount of extracellular EBs in the host medium after one week (for 10% dilution) and after one month (for 0.1% dilution). Briefly, this was done as follows. Culture supernatant was filtered through 1.2 *µ*m syringe filters (Sartorius, Göttingen, Germany) to remove residual host cells. Purified bacterial EBs were then collected by centrifugation (10,000 rpm, 10 min, 4 °C), resuspended in precooled SPG buffer (75 g/l sucrose, 0.52 g/l KH_2PO_4, 1.53 g/l $NaHPO_4 \cdot 7H_2O$, 1.53 g/l $Na_2HPO_4 \cdot 2H_2O$, 0.75 g/l glutamic acid; pH 7.2), homogenized using a 21-gauge injection needle (B. Braun, Melsungen, Germany), and stored at -80 °C in SPG buffer. For quantification of purified EBs, 10 *µ*l cell suspensions in 10 ml Page's amoeba saline (PAS) (0.12 g/l NaCl, 0.004 g/l $MgSO_4 \cdot 7H_2O$, 0.004 g/l $CaCl_2 \cdot 2H_2O$, 0.142 g/l Na2HPO4, 0.136 g/l KH_2PO_4) were filtered onto a polycarbonate membrane with a pore size of 0.2 *µ*m (EMD Millipore, Billerica, MA, USA); cells were stained with 0.1 *µ*g/ml 4',6'-diamidino-2-phenylindole (DAPI) and counted using an epifluorescence microscope (Axioplan 2 imaging; Carl Zeiss, Oberkochen, Germany). The mean number of extracellular EBs released per amoeba per week in the 10% dilution treatment was 16 and the mean number released per month in the 0.1% dilution treatment was 106.

Fluorescence *in situ* hybridization (FISH) using a Cy3-labeled probe (E25-454, sequence GGA TGT TAG CCA GCT CAT; Thermo Fisher Scientific, Waltham, MA, USA) was then performed as described elsewhere (Poppert et al. 2002; Wagner et al. 2003; Sixt et al. 2013) so as to count, on average, how many bacterial particles were present in a fully infected amoebae, which was ~109. Using the number of EBs isolated from both dilution treatments, this equated to an extra 0.15 generations per week in the 10% dilution treatment (16 divided by 109) and 1 generation more per month in the 0.1% dilution treatment (106 divided by 109). These results imply that the main mode of bacterial transmission in our experimental setup is vertical and not horizontal, with the final numbers being: 3.35 generations per weekly transfer in the 10% dilution treatment and 11 generations per monthly transfer in the 0.1% dilution treatment.

Infectivity assays

Protochlamydia EBs were freshly purified from *Acanthamoeba* cultures grown in 175 cm^2 culture flasks, in which EBs had been allowed to accumulate in TSY for 1 week.

Purification of EBs was conducted as described above. Symbiont-free amoebae were harvested right before infection and an aliquot was counted in a Neubauer chamber. The amoeba cell suspension in TSY was seeded in the required amount of wells (1 ml/well) in a 24-well plate at a concentration of 10^5 amoebae/ml. The plate was then incubated at 20 °C for 1 hour to allow the amoebae to attach prior to infection. EBs for the samples to be tested were thawed at 37 °C in a water bath and added in triplicate to wells in the 24-well plate at a multiplicity of infection (MOI) of 30, as this was predetermined to be the best MOI to observe differences in infectivity between samples. The well-plates were then incubated for 15 min at room temperature, followed by centrifugation (1,000 rpm, 15 min, 23 °C) so as to synchronize the infection. The medium was then changed so as to remove any residual EBs from the medium. This was done by gently removing 1 ml TSY per well by pipetting, and then re-adding fresh 1 ml TSY. Wells were then sampled at 6 hours post infection (hpi) and 96 hpi, marking the initial uptake of EBs and the completion of one developmental cycle, respectively. One of the wells with amoebae was treated the same way as the rest but left uninfected, and this served as a negative control to ensure that there was no infection prior to incubation.

FISH using a Cy3-labeled probe (E25-454, sequence GGA TGT TAG CCA GCT CAT; Thermo Fisher Scientific, Waltham, MA, USA) was then performed as described elsewhere (Poppert et al. 2002; Wagner et al. 2003; Sixt et al. 2013) on the sampled time points, and cells were also stained with DAPI for 3 min. Using an epifluorescence microscope, the number of infectious particles in 50 amoebae from each sample were counted and placed into one of three categories: uninfected, slightly infected (1–30 particles) and highly infected (> 30 particles), and this allowed for the infectivity of that particular sample from the assay to be established.

Curing amoeba cultures from infection was performed in 25 cm^2 culture flasks by adding the antibiotic Rifampicin (Sigma-Aldrich Handels GmbH, Vienna, Austria). Initially 6 *µ*l of 100 mg/ml Rifampicin solution in DMSO was added. For monitoring the antibiotic effect, the infection was observed by performing FISH weekly. According to the intensity of the remaining infection, another dose of Rifampicin solution was added to the culture and, if necessary, the dose was increased up to 12 *µ*l of the 100 mg/ml Rifampicin solution. The amoeba culture was considered cured when no symbionts could be observed by FISH using the symbiont specific probe.

DNA extraction and sequencing

Frozen amoeba-chlamydia cultures were thawed in a water bath at 35 °C for 2 min and then pipetted into 25 cm^2 culture flasks (Nalge Nunc International, Rochester, NY, USA) with 10 ml TSY to allow sufficient growth. After one week, these cultures were then upscaled into 175 cm^2 flasks (Biologix Group Ltd, Shandong, China) for another week so as to increase the quantity of EBs that could be isolated. EB purification from the supernatant was done as described above. Isolation of genomic DNA was then carried out on these isolated EBs using a cetyl trimethylammonium bromide-based extraction method (Schmitz-Esser et al. 2010). Library preparation and pool sequencing of symbiont populations were performed at the Vienna BioCenter Core Facilities (VBCF) Next-Generation Sequencing (NGS) Unit (https://www.viennabiocenter.org/facilities/). Illumina Genome Analyser and HiSeq2000 instruments were used to generate paired-end reads of ~125 bases according to standard procedures, achieving the desirable coverage of 300- to 600-fold.

Sequence read processing and mapping

The quality filtering and trimming of the Illumina reads was done by Prinseq-lite (v0.19.5) (Schmieder and Edwards 2011) and Trimmomatic (v0.32) (Bolger et al. 2014) and briefly applied as follows. First, a sliding window of size 10 removed any bases with lower quality than 20 starting from the 3' side by removing the read part containing the low-quality bases. Any reads shorter than 40 nucleotides were also removed. We discarded any low quality read that had an average Phred score below 30. Only read pairs were kept. These reads were mapped against the reference genome sequence of *P. amoebophila* UWE25 (Horn et al. 2004) (NCBI RefSeq accession number NC_005861.2) using the Burrows-Wheeler Aligner (BWA) (v0.7.5a) (Li and Durbin 2009) with standard settings and stored as bam files. For conversions from sam to bam files and from bam to fastq files (as Cortex_var input), we used SAMtools (v0.1.18) (Li et al. 2009) and Picard Tools (v1.92, http://broadinstitute.github.io/picard/).

Variant prediction

We used the VarCap pipeline (https://github.com/ma2o/VarCap) (Zojer et al. 2017) to detect single-nucleotide and structural differences at high resolution and accuracy

based on Illumina reads mapped to the founder ancestral genome sequence of *P. amoebophila*. VarCap is able to detect all relevant prokaryote genome variant types simultaneously with high accuracy. VarCap automatically performs quality filtering, mapping, variant calling and post-filtering of the predicted variants. The output is a Variant Call Format (VCF) file with a detailed description of the variants and their relative frequencies as well as two PDF files, which give a graphical overview of variant coverage and their frequency distribution.

Owing to the different variant calling abilities of the separate tools at low frequencies, various tools were combined so as to increase the sensitivity of the VarCap pipeline (Varscan2, LoFreq-Star, Pindel, Breakdancer, Delly, Cortex_var). Furthermore, a variant call had to be supported by at least two different tools so as to gain precision and robustness. Combining all these selected software tools, all variants except inversions were able to be detected at a minimum read abundance of 2% at 400x total coverage with a minimum of 8 reads per variant, with high sensitivity. Repetitive regions that were larger than the insert site were flagged in order to mark any variants that appeared within these regions for further inspection. This was done so as to avoid false positives (FP) due to reads mapped to repetitive regions. Apart from this, to resolve FP that resulted from the incomplete detection of the true variant type, larger variants were prioritised over smaller ones. Thus, smaller variants were assigned to larger ones when they described a component of the entire variation, such as a large indel at the excision site of translocation.

The variant calling workflow was tested and evaluated using different reference bacterial and archaeal genomes, including *P. amoebophila* (Zojer et al. 2017). Low abundance (2%) variants were detected, with three being confirmed by PCR and Sanger sequencing. VarCap is therefore an ideal pipeline for our purpose and was selected not only due to its thorough evaluation but especially because it already had demonstrated its suitability at detecting variants in our setup.

Estimation of mutation rate

Mutation accumulation (MA) experiments in microbes are designed in such a way as to allow mutations to arise in a neutral manner with minimal section by repeatedly imposing tight population bottlenecks of one or a few cells for many generations (Halligan and Keightley 2009). The spontaneous mutation rate can then be estimated by simply counting the number of genetic changes that would have occurred in the

evolved genome after the given amount of generations. Although we were not able to streak for single colonies on agar medium to accomplish the tight bottleneck, we consider the treatment of 0.1% dilution at 20 °C to be somewhat of an MA experiment. The ambient temperature of the experiment imposed minimal selection on the host-symbiont system, whereas the 1,000-fold dilution reduced the effective population size drastically every month. This allowed around 11 generations of growth between every dilution, which is even fewer generations per passage than MA experiments carried out on *Salmonella typhimurium* (25 generations) and *E. coli* (28 generations) (Lind and Andersson 2008; Lee et al. 2012).

We estimated the mutation rate in two ways, one likely resulting in an overestimated value and thus an upper bound and the other resulting in a much more conservative value. In both cases we considered the three replicates from the 0.1% dilution at 20 °C treatment that were sequenced over multiple time points including the final one at 420 generations. In the first estimation we added up all identified mutations (SNPs and small indels < 4 nt) per replicate across all the sequenced time points. We then calculated the mutation rate by taking the average number of mutations per replicate and dividing by (i) genome size and total number of generations for the per nucleotide rate and; (ii) total number of generations for the per genome rate. This method clearly resulted in an exaggerated mutation rate since these mutations were identified in a population of individuals and not in single clones. Conversely, the calculations for the second estimate only considered those mutations present in the final time point and whose frequency was at least 50% in the population, implying that they were present in the major clones of that specific population.

Statistics

Statistical analysis was performed using R (v3.6.1) (R Development Core Team 2011)**.** The statistical parameters and significance are reported in the figure legends. Data were considered to be statistically significant when $P \leq 0.05$. Comparisons for two groups were calculated by a two-sided Wilcoxon–Mann–Whitney test and comparisons for more than two groups were calculated by a Kruskal-Wallis test followed by Dunn's multiple comparison post-hoc test. Comparisons for paired samples were calculated by a paired sample *t*-test or a Wilcoxon signed-rank test.

The presence of a trend across multiple time points was tested using a Mann-Kendall test.

To determine the specificity of genomic evolution, we followed the calculations carried out by Deatherage and colleagues, only including those mutations that solely impacted one single gene, termed 'qualifying' mutations (Deatherage et al. 2017). These included nonsynonymous SNPs as well as small insertions and deletions (< 4 nt). Intergenic mutations upstream and within 250 bp of the gene start were also considered to be potential promoters affecting the nearest flanking gene (Pilalis et al. 2011). Dice's Coefficient of Similarity, S, was computed for each pair of evolved replicates, where $S = 2\,|X \cap Y| / (|X| + |Y|)$. Here, $|X|$ and $|Y|$ are the sets of genes that have qualifying mutations in two replicates, and $|X \cap Y|$ is the set that possesses mutations in both replicates. The range of this coefficient is from 0, when the two replicates do not share any qualifying mutated genes, to 1, when both replicates have qualifying mutations in exactly the same gene set.

Data availability

All raw sequencing data was deposited in the National Center for Biotechnology Information (NCBI) database under bioproject accession number PRJNA631516.

Results

The evolution experiment

We designed the experiment in such a way so as to be able to look at various aspects of adaptive processes of the *Protochlamydia* symbiont in its *Acanthamoeba* host. We made use of two temperatures (20 °C and 30 °C) and two dilutions (0.1% and 10%) (fig. 1). Starting with an ancestral amoeba culture infected with the strictly intracellular symbionts, four treatments were set up through the different combinations available. The 20 °C treatments served to represent adaptation under natural conditions, whereas the 30 °C served to represent adaptation to elevated temperature. Selection could occur at both dilutions but we chose two very different dilutions in order to vary the importance of genetic drift. In fact, in the 0.1% dilution treatment the impact of genetic drift and the role of chance should be significantly increased relative to the 10% dilution treatment.

The evolution experiment consisted of regular transfers of infected amoeba host cells to fresh media and ran for a total of 38 months, equivalent to 420 (0.1%

dilution) or 510 (10% dilution) bacterial generations (see Materials and Methods section for more information). Phenotype assays and pool sequencing of numerous replicates was performed at different time points throughout the experiment (fig. 1). The ancestral *P. amoebophila* founder strain was used as a reference to identify all variants that arose throughout the experiment (supplementary table S1, Supplementary Material online). 104 samples were sequenced in total across six time points. We sequenced fairly regularly every 5–7 months as we predicted that this allowed sufficient time for a few mutations to arise and be selected for in the different treatments. We sequenced a larger number of replicates from the 30 °C treatments (n = 62) as we were particularly interested in adaptation to this elevated temperature. At both temperatures we analysed more replicates from the 10% dilution treatment as the impact of genetic drift was much less pronounced in these treatments, and we were keen to identify as many variants as possible that were potentially beneficial and therefore increased in frequency in the population. This greater sequencing effort would allow us to identify any cases of parallelism among replicates.

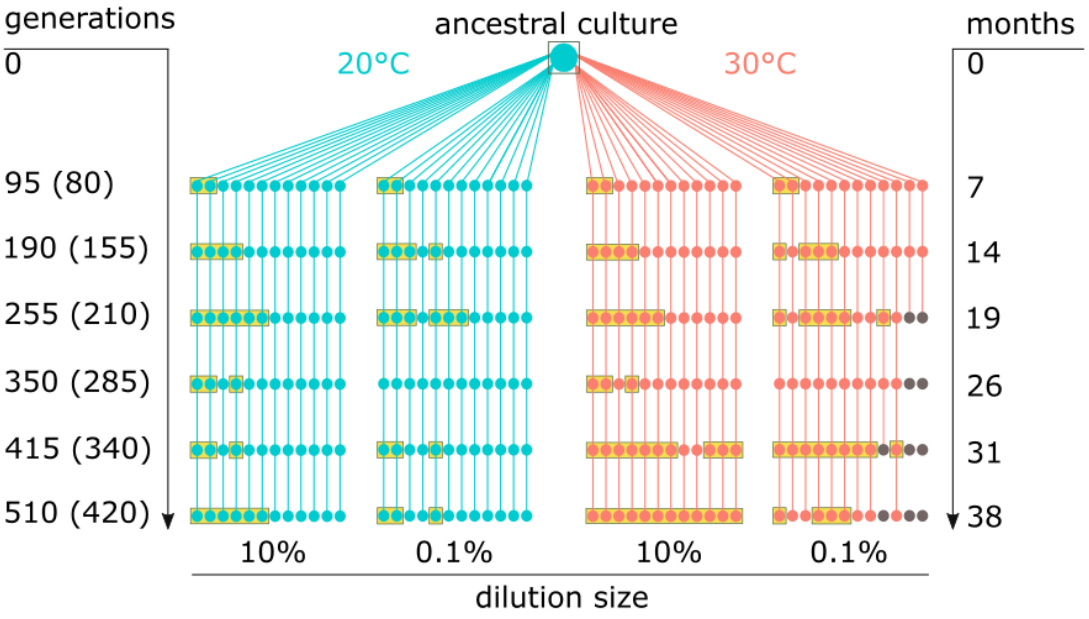

FIG. 1. Overview of experimental setup and sequenced samples in the evolution experiment. Pooled whole genome re-sequencing of symbiont populations was performed on numerous replicates (n = 104) from all four treatments. Sequenced *Protochlamydia amoebophila* populations are shown in boxes. Grey circles represent populations that died out during the experiment. The relevant numbers of generations for the 0.1% dilution treatments are shown in brackets.

Temperature-dependent attenuation of infectivity

The fitness of strictly intracellular bacteria is largely determined by their ability to establish and maintain infections in their host cell population. We thus performed infection assays to determine whether the infectivity of the evolved symbionts was different from that of the ancestral founder population. This was carried out by infecting naive (symbiont-free) ancestral amoebae with EBs isolated from a specific treatment at a particular time point and measuring the mean percent of highly infected amoebae 96 hours post infection (hpi), equivalent to the duration of one full developmental cycle of the symbiont (König et al. 2017). The infection assays were performed at both ambient (20 °C) and elevated (30 °C) temperatures.

Overall we observed that the infectivity remained fairly constant throughout the evolution experiment at either assay temperature for the symbionts isolated from the 20 °C treatment replicates, but the infectivity was higher in assays performed at the ambient temperature from which the symbionts were isolated originally (supplementary fig. S1*A* and *B*, Supplementary Material online). Symbionts from the 30 °C treatment replicates displayed a lower infectivity at either assay temperature than those symbionts from the 20 °C replicates (supplementary fig. S1, Supplementary Material online). As the experiment progressed, the infectivity of the symbionts from the 30 °C replicates significantly decreased further, especially when the assays were performed at ambient temperature (supplementary fig. S1*C* and *D*, Supplementary Material online). This implies that as the experiment progressed, the symbionts evolved at the elevated temperature developed an attenuated infectivity.

We performed the most comprehensive analysis with symbionts from all four treatments obtained after 31 months (340 generations for 0.1% dilution treatments and 415 generations for 10% dilution treatments). This shows that at ambient temperature the infectivity of symbionts evolved at 20 °C did not significantly change at either dilution when compared to ancestral EBs (fig. 2*A*). Confirming our previous analysis, this is not the case with EBs evolved at the higher temperature, where a pronounced and significant decrease in infectivity was measured at both dilutions (fig. 2*A*). When performing the infectivity assays at 30 °C, the infectivity of the symbionts was significantly lower than at 20 °C for the ancestor and the populations that evolved at 20 °C (fig. 2*B*). Although the infectivity of symbionts from the 30 °C treatment was the lowest at this temperature, it was still significantly higher than the infectivity of those same symbionts when the assays were performed at 20 °C (*P* =

0.020; fig. 2*B*). This suggests that the chlamydiae maintained for a long period at the elevated temperature did in fact adapt to these environmental conditions.

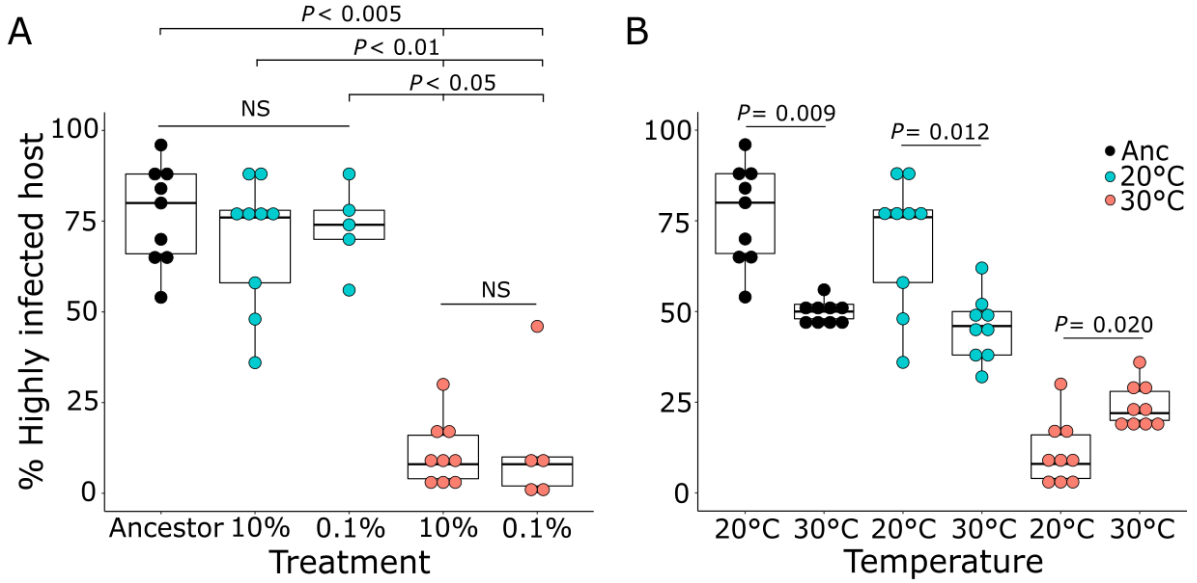

FIG. 2. Infectivity of ancestral and evolved symbionts. Naive amoebae (*n* = 9 ancestor and 10% dilution, *n* = 5 0.1% dilution) at a concentration of 10^5 cells per ml were infected with *P. amoebophila* symbionts isolated from the ancestor and evolved populations after 340 generations (0.1% dilution) or 415 generations (10% dilution) with a multiplicity of infection (MOI) of 30. The percent of highly infected amoebae was assessed at 96 hpi using fluorescence *in situ* hybridization. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. **(*A*)** Infectivity assays performed at 20 °C. The infectivity among the different treatments was compared using a Kruskal-Wallis test ($\chi^2 = 25.494$, $p < 0.0005$) followed by Dunn's multiple comparison post-hoc test. **(*B*)** Comparison of infectivity assays performed at 20 °C and 30 °C for the ancestor and both 10% dilution treatments. The infectivity between paired samples at different temperatures was compared using a Wilcoxon signed-rank test.

Since these experiments were carried out using ancestral amoeba host cells that were maintained in the laboratory environment at 20 °C, we wanted to account for this potential bias relating to the host used. To this end, we also carried out a subset of the infectivity assays using uninfected amoebae that had evolved at 30 °C for the duration of the experiment in order to investigate the host effect on infectivity. There was no observed difference in infectivity for any of the symbionts when performing the infectivity assays using amoebae adapted to the elevated temperature (supplementary fig. S2, Supplementary Material online). We also cured

the amoebae from two replicates in the 30 °C treatment from their chlamydial symbionts using rifampicin. These symbiont-free amoebae were then re-infected, at both 20 °C and 30 °C, with co-evolved EBs that had been isolated prior to the curing procedure. Once again, there was no change in infectivity for these 30 °C evolved EBs when re-infecting their co-evolved hosts, demonstrating that the reduction in infectivity was not due to the introduction to a novel host, but was a function of the symbionts themselves (supplementary fig. S3, Supplementary Material online). Finally, in order to assess the stability of this loss of infectivity, we took populations from three replicates in the 30 °C treatment and maintained them at 20 °C for eight weeks. After performing infectivity assays, it was evident that there was no significant increase in infectivity of these EBs at the lower temperature as compared to their infectivity after isolation from the 30 °C treatment (data not shown). This implied that the loss of infectivity in the 30 °C populations was a stable phenotype.

Dynamics of genome evolution

Next we performed pooled whole genome re-sequencing of in total 104 symbiont populations from diverse time points and replicates (fig. 1). We identified all the variants that arose throughout the evolution experiment in all the replicates across all treatments, and this amounted to 380 mutations (supplementary table S1, Supplementary Material online). The majority of the variants were single nucleotide polymorphisms (SNPs, 83.4%) as opposed to small insertions/deletions (indels < 4 nt, 16.6%), and from these SNPs, a much larger proportion was identified as being in the protein-coding region (83.0%) compared to the intergenic region (17.0%). Roughly three quarters of the SNPs in the protein-coding region were nonsynonymous mutations (74.9%), with a quarter of the SNPs (25.1%) being synonymous and not causing any alteration in the encoded amino acid. Distinguishing between synonymous and nonsynonymous mutations is of great significance because selection has different impacts on these two types of mutations (Yang and Nielsen 1998). Whereas synonymous mutations are generally silent and have no detectable effect (but see Agashe et al. 2016 and Lebeuf-Taylor et al. 2019), nonsynonymous mutations are typically harmful, are selected against, and will be eliminated by purifying selection (Jia et al. 2003). In those rare cases when nonsynonymous mutations confer a selective advantage, they are favoured and maintained by positive selection. Our results would therefore suggest that many of

the SNPs that arose were beneficial and were selected to increase in frequency in the population.

Six replicates from the 10% dilution treatment, three from each temperature, were sequenced at every time point from 190 generations onwards (fig. 1).These are referred to here as ‘core replicates’ and allow for more comprehensive analyses of evolutionary changes during the experiment. In total we identified 121 SNPs and 19 small indels (< 4 nt) across all six core replicates. No apparent mutational hotspots were observed across the 2.41 MB genome of *P. amoebophila* (fig. 3). Fewer mutations (*n* = 36) were identified in the core replicates at ambient temperature and the majority of these were SNPs, with only one indel present (fig. 3*A*). The core replicates at the elevated temperature had more than double the number of SNPs (*n* = 86) at ambient temperature, with 18 indels also being identified (fig. 3*B*). Overall, synonymous and intergenic mutations each amount to around 20% of all SNPs, with the majority (60%) of the SNPs being nonsynonymous. A majority of the variants identified in the 20 °C replicates were low frequency, and after 510 generations only eight had a frequency above 15%, with three of those above 60% (fig. 3*A*). On the other hand, the variants identified in the 30 °C replicates reached higher frequencies. After 510 generations 23 of these SNPs had a frequency above 15%, with ten of those above 60% (fig. 3*B*).

FIG. 3. Emergence of mutations across the symbiont genome during 510 generations. Variants in the *P. amoebophila* genome during long-term evolution in the core replicates at

(*A*) ambient and (*B*) elevated temperature are shown as circular plots. The five grey rings, from inner to outer, show mutations present in genomes sampled at 190, 255, 350, 415 and 510 generations. SNP mutations are separated according to whether they are nonsynonymous (red), synonymous (blue) or intergenic (yellow). Indels are separated according to whether they are deletions (purple) or insertions (orange). The frequencies of the mutations present at 510 generations are also shown in the outermost green ring. Genome visualization was carried out with Circos (Krzywinski et al. 2009). Molecular details for all mutations are shown in supplementary table S1, Supplementary Material online.

We estimated the spontaneous mutation rate of the *P. amoebophila* symbiont using two methods, which resulted in an upper bound value and a much more conservative value (see Materials and Methods section for more information; table 1). We chose to do this since we studied populations of individuals in our system and not single clones; therefore, these two estimates should provide the range in which the actual spontaneous mutation rate lies. We made use of the three replicates that evolved at 20 °C with a monthly 1,000-fold dilution as this treatment resembles a mutation accumulation (MA) experiment in which all mutations should accumulate at the rates at which they happen, regardless of fitness effects, with the exception of lethal and highly deleterious mutations (Barrick and Lenski 2013). This resulted in an estimated spontaneous mutation rate between 1.11 x 10^{-8} and 9.76 x 10^{-10} mutations per nucleotide per generation (table 1).

Table 1. Estimated spontaneous mutation rates of *Protochlamydia amoebophila*.

Method	No. of SNPs	No. of indels*	No. of replicates	Mutations per line	Mutation rate per nucleotide (x 10^{-8})†‡	Std dev (x 10^{-9})	Mutation rate per genome (x 10^{-2})‡	Std dev (x 10^{-2})
Upper bound	33	1	3	11.33	1.106	8.417	2.698	2.053
Conservative	3	0	3	1.00	0.098	0	0.238	0

*Insertion or deletion of < 4 nt.
†Genome = 2,438,912 nt.
‡Generations per replicate = 420 gen.

We noted that one of the 10% dilution replicates at 30 °C (replicate 4; supplementary table S1, Supplementary Material online) evolved a moderately elevated mutation phenotype at some point between 190 and 255 generations. At its first sequenced time point (190 generations) only two mutations in total had been identified. However, the number of mutations (SNPs and small indels < 4 nt)

identified in this replicate increased around 4-fold at the next sequenced time point and remained elevated when compared to the other replicates of the same treatment throughout the remainder of the experiment (table 2). Five variants were identified in this replicate that arose by 255 generations and subsequently remained in the population throughout the course of the evolution experiment (table 2). The observed elevated number of mutations in this symbiont population was most likely the result of a nonsynonymous SNP located in the *ung* gene that codes for the enzyme uracil-DNA glycosylase. This enzyme catalyzes the first step in a repair pathway for uracil-containing DNA in *E. coli* and in many other organisms (Lindahl 1979), and *E. coli ung* mutants have previously been shown to display elevated mutation rates (Duncan and Weiss 1982).

Table 2. Evolution of an elevated mutation phenotype. The persistent variants and the elevated number of mutations in replicate 4 of the 10% dilution 30 °C treatment are shown.

Variant type	Variant effect	Gene name	RefSeq description	255 gen (%)	350 gen (%)	415 gen (%)	510 gen (%)
SNP	nonsynonymous	NA	hypothetical protein	13.7	28.7	39.0	34.9
SNP	nonsynonymous	rpsA	30S ribosomal protein S1	27.0	74.8	59.4	66.1
SNP	nonsynonymous	rumA	RNA methyltransferase	19.0	24.7	38.7	37.5
SNP	nonsynonymous	ung	uracil-DNA glycosylase	6.4	23.0	42.0	37.4
INS	frameshift	NA	hypothetical protein	3.3	20.3	25.2	30.6
Number of mutations			***Mean of other replicates***	5.6	6.5	10.5	9.3
			Replicate 4	24	19	24	35
			Fold-increase	4.3	2.9	2.3	3.8

Evidence for positive selection

By identifying all the mutations that arose during the evolution experiment, we were able to compare the mean number of accumulated synonymous and nonsynonymous substitutions in the symbiont populations for each treatment (fig. 4). An overrepresentation of nonsynonymous substitutions indicates that the main driving force for mutations was positive selection as those variants would have provided an adaptive advantage causing an increase in their frequency in the population (Barrick et al. 2009). The lower temperature should not have been stressful for either host or symbiont as 20 °C represents standard laboratory conditions used for culture maintenance. Yet, we observed that there was a

statistically significant increase in the number of nonsynonymous substitutions relative to the synonymous ones in the 10% dilution treatment at the final time point (*P* = 0.0098, fig. 4*A*). This suggests that a certain level of adaptation to the lab conditions continued to occur, irrespective of the ambient temperature they were maintained in. In the higher dilution at the ambient temperature we observed fewer mutations accumulate, with a very similar proportion of synonymous and nonsynonymous substitutions (fig. 4*B*).

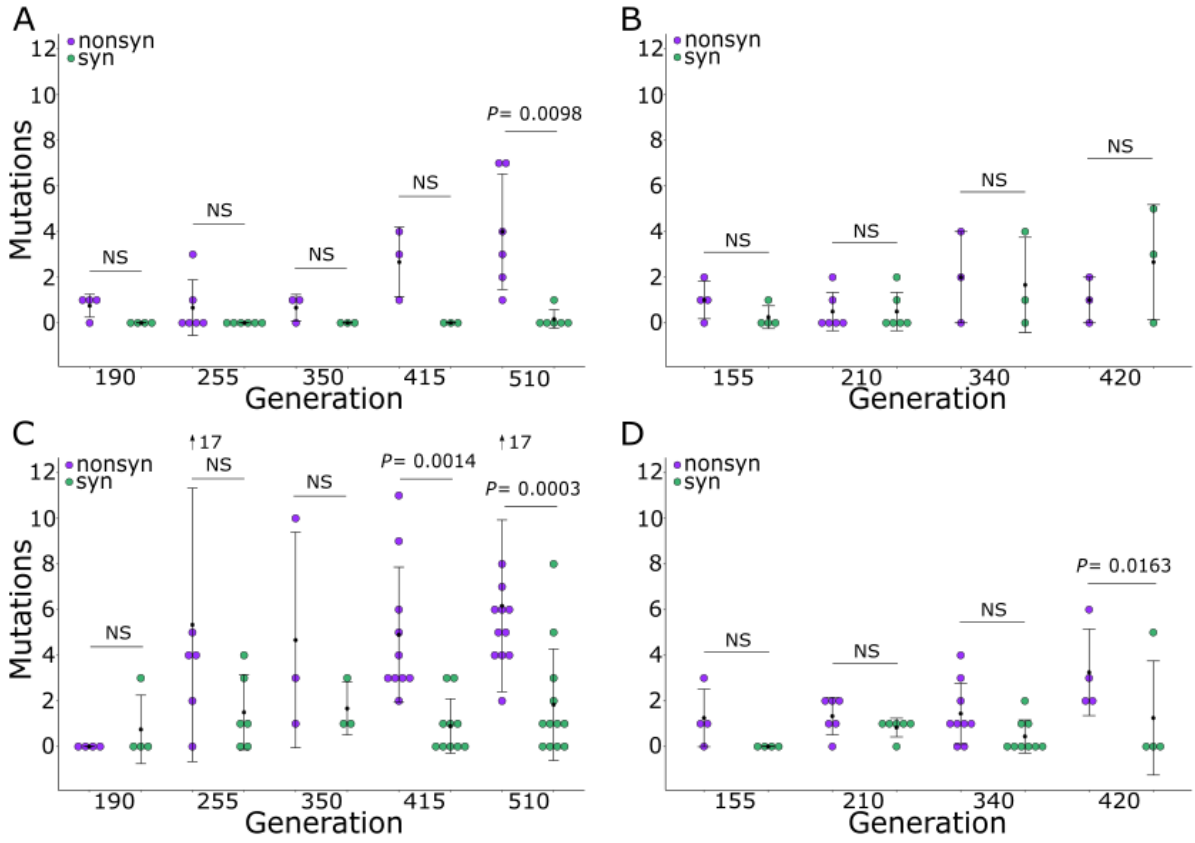

FIG. 4. Accumulation of nonsynonymous substitutions. The number of nonsynonymous and synonymous substitutions throughout the experiment are shown for the (*A*) 10% dilution at 20 °C treatment; (*B*) 0.1% dilution at 20 °C treatment; (*C*) 10% dilution at 30 °C treatment; and (*D*) 0.1% dilution at 30 °C treatment. Data represent mean ± standard deviation of the mean. Symbols denote numbers for individual replicates. The two arrows in (*C*) represent two outliers. Synonymous and nonsynonymous mutations per time point were compared using a paired sample *t*-test (significance indicated above pairs).

The elevated temperature was expected to create a more challenging environment for the chlamydial symbionts as both host and bacteria are used to thriving at ambient temperatures in their natural environment, well below 30 °C. In fact, we saw the largest accumulation of mutations in the 10% dilution treatment at

30 °C (fig. 4*C*). The last two time points revealed a significant increase in the number of nonsynonymous substitutions relative to the synonymous ones ($P = 0.0014$ at 415 gen; $P = 0.0003$ at 510 gen). These nonsynonymous substitutions were likely beneficial, allowing them to increase in frequency in the symbiont populations. Furthermore, we also observed a similar but less pronounced increase at the last time point in the 0.1% dilution treatment ($P = 0.0163$) (fig. 4*D*), despite the impact of genetic drift predicted to be stronger due to the bottleneck effect of the high dilution during passaging.

Apart from comparing SNPs across the different treatments, we wanted to estimate the proportion of putatively beneficial changes for small indels (< 4 nt). We did this by comparing the number of small indels that arose in the 0.1% dilution treatment at 20 °C, which resembles a mutation accumulation experiment, with the other three treatments (table 3). We observed that small indels were very rare in all treatments apart from the 10% dilution at 30 °C treatment (table 3). The replicates from this latter treatment accumulated the highest number of small indels, on average, at each one of the four time points. In particular, we identified a mean greater than two small indels per replicate for each of the last two sequenced time points (table 3). Small indels were therefore overrepresented in the 10% dilution at 30 °C treatment relative to the MA experiment, implying that many of these mutations were adaptive during the experiment and selected to increase in frequency in the population.

Table 3. Small indels identified across all treatments. The relevant number of generations for the 0.1% dilution treatments are shown in brackets.

Treatment	Generations							
	190 (155)		255 (210)		415 (340)		510 (420)	
	Mean	*Std Dev*	*Mean*	*Std Dev*	*Mean*	*Std Dev*	*Mean*	*Std Dev*
0.1% dilution 20 °C	0	0	0	0	0.33	0.58	0	0
10% dilution 20 °C	0	0	0	0	0	0	0.17	0.41
0.1% dilution 30 °C	0.25	0.50	0.17	0.41	0.11	0.33	0.25	0.50
10% dilution 30 °C	0.50	0.58	1.17	0.98	3.80	2.39	2.25	1.82

We next analysed the trajectories of mutations over sequential time points in the core replicates (fig. 5). Assuming that mostly beneficial mutations are selected for in all but the mutation accumulation experiment, this analysis should shed light on the mode and strength of selection in the different treatments and the fitness effects of the mutations observed (fig. 5). At first glance it is evident that the 30 °C replicates accumulated more mutations throughout the experiment compared to the ambient temperature. More importantly, we can identify those mutations that persisted over multiple time points, which would indicate that they likely provide a benefit as they were being maintained in the population. In fact, by the final time point, the amount of mutations identified at 30 °C that had already been present earlier in the experiment (n = 26) were more than double the amount at 20 °C (n = 10). Nine mutations were already present in the 30 °C treatment replicates from 255 generations, whereas none of the mutations in the 20 °C treatment replicates persisted for this long. At 20 °C, mutations tended to arise at one time point and were subsequently replaced by mutations at the next time point, therefore not really providing any long-term benefit. Finally, overall, mutations reached higher frequencies at the elevated temperature with a median frequency of 20% at the final sequenced time point, as opposed to 6% at the ambient temperature. In fact, at 510 generations only three mutations reached frequencies higher than 60% at 20 °C, whereas we identified ten such mutations at 30 °C. This provides further evidence for positive selection particularly at the elevated temperature and suggests the presence of a number of beneficial mutations in the respective evolved symbiont populations.

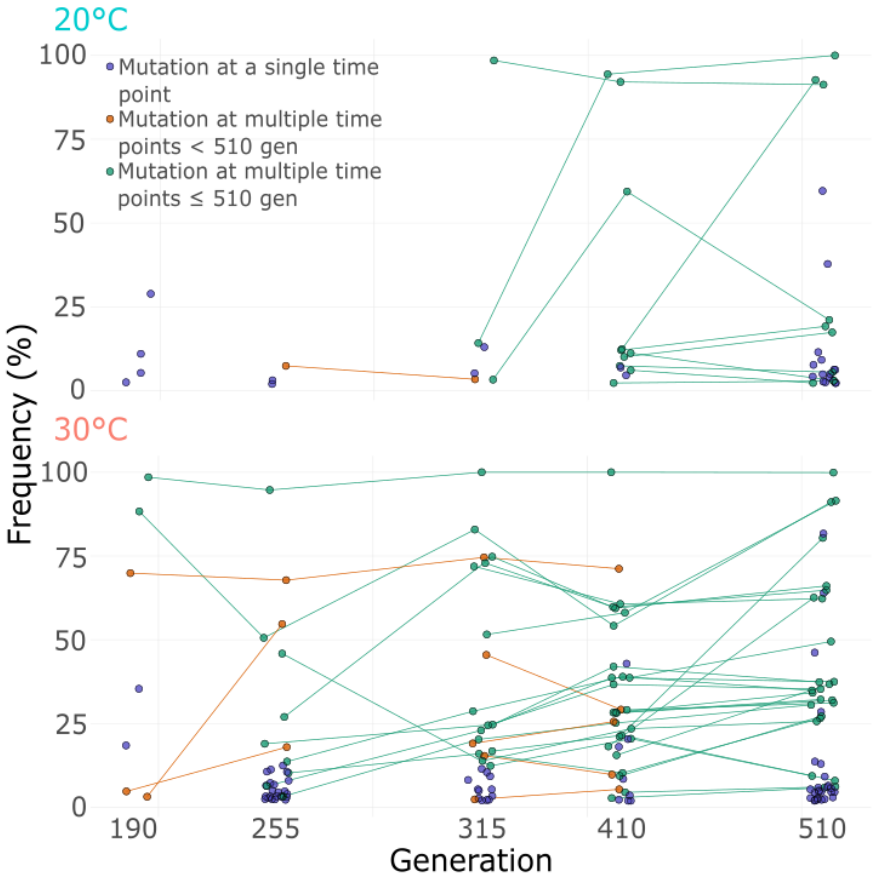

FIG. 5. Mutational trajectories of symbiont populations at ambient and elevated temperature. Trajectories of novel mutations throughout the experiment in the core replicates at 20 °C (top) and 30 °C (bottom) are shown. Following all the novel mutations that arose in the core replicates, we observe that at 30 °C (i) more mutations are present; (ii) mutations are more likely to persist over multiple time points; and (iii) mutations tend to reach higher frequencies.

Since the novel mutations that persisted are generally expected to provide an adaptive advantage, we took a closer quantitative look at the differences in the number of persistent mutations in the core replicates - here defined as mutations (SNPs and small indels < 4 nt) identified in subsequent time points at a frequency of at least 5% (fig. 6). Few persistent mutations were observed at both temperatures for the first 190 generations. However, as the experiment progressed, those mutations increased in number. This was particularly prevalent at 30 °C, where by the last time point the mean number of persistent mutations was nearly four times greater than the mean number at 20 °C (8.67 vs. 2.33). This confirmed our previous observation (fig. 5) that mutations at the elevated temperature were more likely to persist over multiple time points. In fact, there was a statistically significant increasing trend of

persistent mutations over time at 30 °C ($P < 0.05$, Mann-Kendall test) that was not seen at 20 °C (fig. 6).

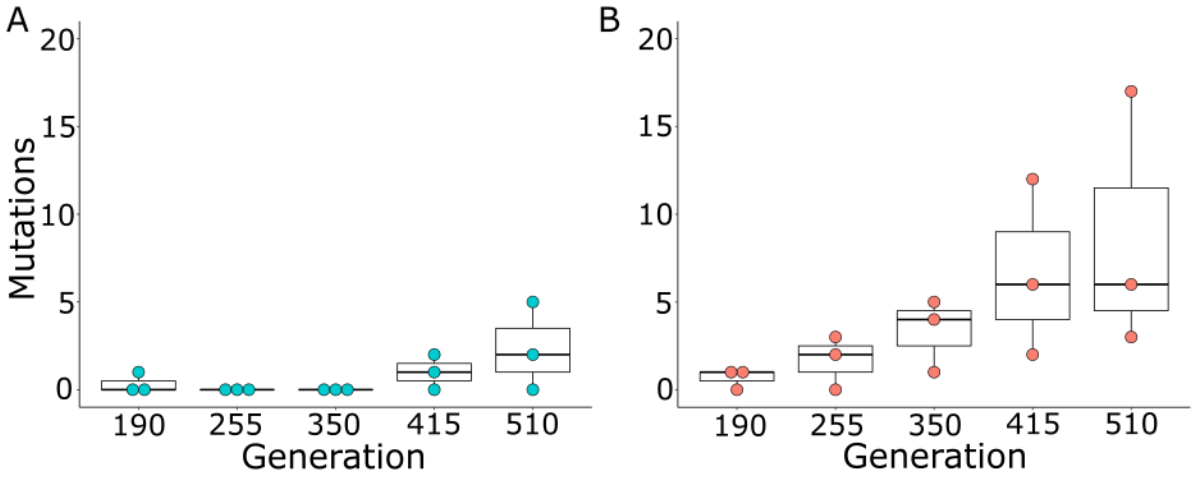

FIG. 6. Temperature-dependent differences in number of persistent and likely adaptive mutations. The number of persistent mutations (SNPs and small indels with a frequency > 5%) throughout the experiment for core replicates at (*A*) 20 °C; and (*B*) 30 °C is shown. Whereas the number of persistent mutations in the replicates at 20 °C remained low at all the time points, persistent mutations steadily increased at each time point in the replicates at 30 °C ($P < 0.05$, Mann-Kendall test).

Convergence and divergence among replicate populations

Our experimental set-up also allowed us to study the specificity of genomic evolution i.e. the extent of parallel evolution during temperature adaptation of the symbiont populations. To this end, we compared mutations identified at the final sequenced time point between replicates within the same treatment (temperature and dilution) and across the different treatments. We only included those mutations that solely impacted one single gene, termed 'qualifying' mutations (Deatherage et al. 2017) (see Materials and Methods section for more information). A total of 113 qualifying mutations was identified to fit our criteria for the analysis. Dice's Coefficient of Similarity, *S*, was then computed for each pair of evolved replicate populations.

Across all replicates from the four treatments, the grand mean similarity, S_m, is 0.02. There were no mutated genes in common at all in the 0.01% dilution treatments at neither temperature. This is not unexpected as the tight bottleneck would have eliminated much variation and the mutated genes observed would have mainly been present due to chance survival. The mean within-treatment similarity, S_w, for the 10% dilution treatments is 0.06 and 0.05 at 20 °C and 30 °C, respectively.

The mean between-treatment similarity, S_b, between these two treatments is 0.02. Replicates from the same treatment were therefore more similar to each other with respect to their mutations than replicates from a different treatment, indicating an adaptation to their particular experimental conditions. This confirmed our previous observations regarding positive selection described above. 17 genes were affected by at least two qualifying mutations across the different treatments (supplementary table S2, Supplementary Material online). Seven genes carried mutations at both temperatures, whereas ten genes were affected by mutations exclusively at 30 °C. No genes were solely affected by qualifying mutations across multiple replicates at the lower temperature.

We next asked, which gene functions were affected by mutations in our experiment and thus putatively associated with temperature adaptation. We therefore separated all the mutations identified in the final sequenced time point across both temperatures into functional groups based on the eggNOG database (Huerta-Cepas et al. 2019) (fig. 7 and supplementary table S3, Supplementary Material online). Genes without a functional prediction (i. e. no match in eggNOG or assignment to categories R and S) were not taken into account. We also only considered replicates from the 10% dilution treatment as we were interested in any mutations resulting in adaptive processes, and wanted to omit mutations present as a result of stochastic events. This analysis showed that the affected functional categories differed considerably between the two temperature regimes. Mutations at 20 °C affected eight functional categories, which is half as many as those affected by mutations at 30 °C (n = 16). The only functional category with a greater mean number of mutations per replicate at ambient temperature was carbohydrate transport and metabolism (fig. 7*A*). On the other hand, there were 14 functional categories with a higher mean number of mutations per replicate at the elevated temperature. In particular, mutations were identified in eight functional categories related to metabolism at 30 °C as opposed to only two at 20 °C (fig. 7*A*). Apart from this, the mean number of mutations in three other functional categories were elevated at 30 °C when compared to 20 °C: (i) J - translation, ribosomal structure, and biogenesis; (ii) O - posttranslational modification, protein turnover, and chaperones; and (iii) V - Defense mechanisms (fig. 7).

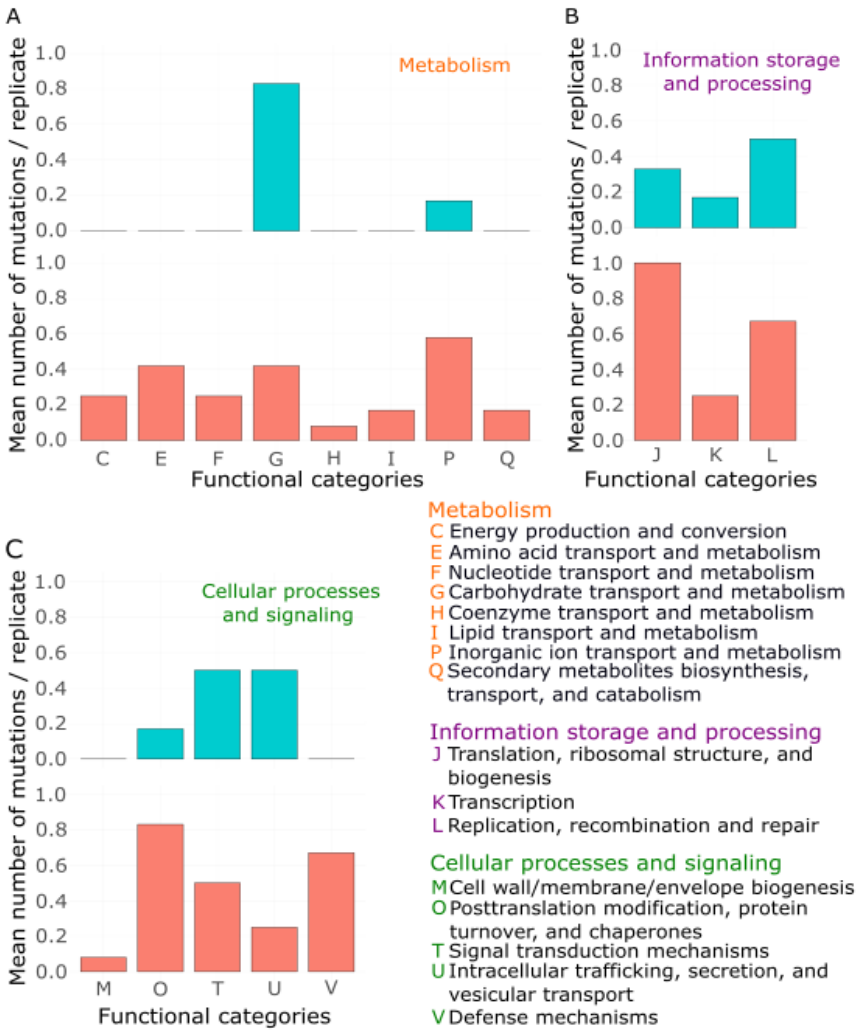

FIG. 7. Cellular functions affected by mutations. The number of mutations assigned to cellular functions including (*A*) metabolism; (*B*) information storage and processing; and (*C*) cellular processes and signaling after 510 generations in the 10% dilution populations at 20 °C (blue bars) and elevated temperature (red bars) is shown. The affected functional categories differed considerably between the two temperature regimes. Mutations affected eight functional categories at 20 °C and 16 functional categories at 30 °C (supplementary table S3, Supplementary Material online).

Discussion

Our evolution experiment using an obligate intracellular symbiont of unicellular eukaryotes was designed to provide insights into basic features of molecular evolution and temperature adaptation in host-dependent microbes. Replicate *Protochlamydia amoebophila* populations were evolved for up to 510 generations in their *Acanthamoeba* hosts under four different combinations of temperature and dilution during passaging. We compared genomic adaptation across the different treatments and also determined whether any phenotype changes occurred

throughout the experiment, in particular to identify how the bacterial symbionts adapted to survive at the higher temperature of 30 °C. This condition should be considered quite a stressful environment for the symbiont and the host since they naturally dwell at lower temperatures.

Attenuated infectivity maintains fitness of microbe-host association at elevated temperature

We measured infectivity of the symbionts as a proxy for bacterial fitness and observed that the infectivity remained fairly constant throughout the evolution experiment for the symbionts from the ambient temperature treatment (fig. 2*A* and supplementary fig. S1*A* and *B*, Supplementary Material online). In contrast, symbionts from the elevated temperature treatment displayed an increasingly lower infectivity as the evolution experiment progressed (fig. 2*A* and supplementary fig. S1*C* and *D*, Supplementary Material online). Importantly, near the end of the observation period the infectivity of the 30 °C evolved symbionts was significantly higher when the assays were performed at the elevated temperature rather than at 20 °C (P = 0.020; fig. 2*B*). Therefore, although the symbionts isolated from the 30 °C treatment developed an attenuated infectivity at either assay temperature as the experiment progressed, these chlamydiae did in fact adapt to the elevated temperature.

Overall temperature adaptation of the symbionts was facilitated through attenuated infectivity. One potential scenario for this is that this phenotype evolved as a trade-off to reduce host cell burden. In fact, we did notice that a proportion of *Acanthamoeba* cells in the cultures maintained at the higher temperature were not as highly infected as amoeba maintained at 20 °C (data not shown). It is thus conceivable that at the elevated temperature there is a lower maximum bacterial load of the *Acanthamoeba* host cells. Therefore, the reduction in infectivity of the 30 °C evolved symbionts would have been a trade-off to reduce host cell stress and compensate for the lower maximum bacterial load per amoeba at the elevated temperature. Ultimately, this would have been beneficial as it would have stabilized the fitness of the whole system. Similar findings have been shown in other endosymbiotic bacteria infecting drosophilid flies, relating to the male-killing phenotype. This trait is characterized by the production of all, or nearly all, female

broods related to low egg hatch rates, and is caused by *Wolbachia* in *Drosophila bifasciata* females. Parents exposed to elevated temperature resulted in a decrease in bacterial density in eggs as well as a decline in penetrance of the trait, which suggested that a threshold density of *Wolbachia* within embryos is required to result in the male-killing phenotype (Hurst et al. 2000). Another investigation on the effect of high temperature through successive host generations on native and nonnative *Drosophila* host species also revealed a reduction in the male-killing phenotype and infection with the symbiont *Spiroplasma* (Anbutsu et al. 2008). Bacterial infection density was significantly suppressed by the elevated temperature in the nonnative *Drosophila* host, but had minimal effect in the native host.

Increased genetic diversity due to structured populations of host-dependent microbes

Pooled whole genome re-sequencing served to study evolutionary processes during our experiment at the molecular level. The majority of mutations identified across all treatments were SNPs (84.3%) as opposed to small indels < 4 nt (15.7%), and 81.9% of the SNPs were located in protein-coding regions. When we took a closer look at the six core replicates, far fewer mutations were identified in the 20 °C evolved populations (n = 36) than in the 30 °C evolved populations (n = 104), which suggests a higher level of genomic adaptation at the elevated temperature (fig. 3). No clear mutational hotspots were identified across the symbiont genome in either treatment (fig. 3). Overall, mutations in the core replicates increased as the evolution experiment progressed, with a greater number of mutations being identified later in the evolution experiment (see supplementary table S1, Supplementary Material online). This may seem unexpected at first glance as a high initial number of mutations and a subsequent decline in substitution rates has been reported in other microbial experimental evolution studies (Conrad et al. 2009; Lee and Palsson 2010). However, this observation might be a consequence of the obligate intracellular lifestyle and the developmental cycle of the symbionts.

First, bacterial and host cell cycles are tightly linked. The symbionts are predominantly transmitted vertically during host cell replication and are thus dependent on binary fission of their amoeba host cells (see Materials and Methods section for more information). As the host cell cycle is slow, the 38 months our evolution experiment was carried out only equated to 510 generations. To put this

into perspective, an evolution experiment with *E. coli* over 60 days comprised around 1,100 generations (Fong et al. 2005). Our evolution experiment is therefore very much in its infancy where selection is still strong and saturation with new mutations may not have occurred yet.

Second, the strictly host-dependent nature and the primary mode of vertical transmission of our symbionts gives rise to a structured environment. The bacterial populations are divided in many subpopulations—within single amoebae—whose fate is each also dependent on individual host fitness. This could result in the maintenance of increased genetic diversity. In fact, the fixation time required for a mutation increases in structured populations and this, in turn, will increase the chance that another beneficial mutation will appear and compete with the first one (Gordo and Campos 2006). This pronounced competition via clonal interference (Gerrish and Lenski 1998) leads to a higher number of segregating mutations in the population, i.e. higher genetic diversity. An increased variation in structured populations has previously been observed in experiments carried out on free-living bacteria such as *Ralstonia* sp., *Pseudomonas fluorescens,* or *E. coli* when comparing populations evolved in structured and homogeneous habitats (Korona et al. 1994; Rainey and Travisano 1998; Habets et al. 2006). In our structured host-microbe system, a mutation that arises during RB replication must compete with other RBs inside the host and eventually needs to be released as an EB. The EB carrying the mutation must then compete against other EBs to enter a new host, and finally outcompete the chlamydiae already residing within that new host. This creates multiple levels of selection, and an adaptive advantage at one stage may face a trade-off by being at a disadvantage elsewhere. Furthermore, the different subpopulations will each fix beneficial mutations conferring small benefits since they do not have access to the rare beneficial mutations of large effect like in larger, well-mixed populations (Orr 1998; Burch and Chao 1999; Miralles et al. 1999). As a consequence, the symbiont populations would have few mutations that actually become fixed, with low frequency mutations co-existing simultaneously being more common (fig. 5).

Although most of the analysis in Lenski's LTEE was carried out after sequencing clones, they also identified mutational variants in mixed bacterial population samples (Barrick and Lenski 2009). We can roughly compare the number of SNPs identified at the final sequenced time point (510 gen) of our 10% dilution

symbiont populations with their first sequenced time point (2,000 gen). They identified seven substitutions in total, whereas our populations at 20 °C and 30 °C had an average of 5.3 and 9.1 substitutions per replicate, respectively. If we were to extrapolate the number of substitutions to 2,000 generations, it would appear that the symbiont populations at 20 °C and 30 °C have a 3-fold and 5-fold increase in the number of mutations relative to the LTEE, respectively. This elevated number of mutations in our symbiont populations may very well be a result of the increased structure of our host-dependent system relative to the environmentally homogeneous habitat provided in the LTEE (Barrick and Lenski 2009), where mutations can outcompete each other more easily.

Bottleneck effect and spontaneous mutation rate

Natural symbiont populations experience frequent bottlenecks during transmission and during shifts in environmental conditions. To assess the impact of a strong population bottleneck on temperature adaptation we included the 0.1% dilution treatment in our evolution experiment. A sudden reduction in population size causes random genetic drift to have a much greater role in determining gene frequency changes through stochastic effects (Nei et al. 1975). Despite there being an identical 1,000-fold dilution at ambient and elevated temperatures, we did observe differences in genome dynamics between the two treatments with respect to the ratio of nonsynonymous mutations to synonymous mutations, which was almost three times higher at the elevated compared to the ambient temperature (fig. 4*B* and *D*). This suggests that at 30 °C mutations were likely adaptive and selection may have still played an important role even under the bottleneck condition used in our experiment.

The 0.1% dilution treatment at ambient temperature also served as a mutation accumulation experiment, and we used this minimal selection condition to estimate the spontaneous mutation rate of our bacteria employing two methods (see Materials and Methods section for more information; table 1). The typical rate of spontaneous mutations obtained from such experiments with bacteria is of the order of 10^{-10} to 10^{-9} per nucleotide per generation (Lind and Andersson 2008; Lee et al. 2012; Sung et al. 2012). The rate calculated by Drake for microbes is 3-4 x 10^{-3} per genome per generation (Drake 1991; Drake 2009). With 1.1 x 10^{-8} mutations per nucleotide per generation and 2.7 x 10^{-2} per genome per generation our upper bound estimate represents an overestimate of the actual mutation rate (table 1), at least one order of

magnitude greater than previous estimates on bacteria. This was expected since we included every single mutation identified throughout the experiment in our calculations. Our more conservative estimate, however, falls within the range determined for other bacteria (table 1). For this estimate we only considered mutations at 420 generations whose frequency in the population was at least 50% and, thus, present in the major clone of that specific population. The more conservative value lies on the lower end of the previously calculated mutation rates and, in fact, the substitution rate of beneficial mutations is expected to be lower in a structured environment relative to a homogeneous population, where purging via selective sweeps should be a relatively rapid process (Gordo and Campos 2006) Overall, our estimate of the substitution rate for the chlamydial symbiont *P. amoebophila* is consistent with values inferred from long-term laboratory propagation of *C. trachomatis* (Borges et al. 2013) or calculated using a phylogenetic approach for *C. trachomatis* and other bacteria (Hadfield et al. 2017).

Positive selection and adaptive evolution at elevated temperature

Several lines of evidence indicate positive selection at 30 °C. This includes a significantly higher amount of nonsynonymous mutations relative to synonymous mutations at the last two sequenced time points (fig. 4*C*). Apart from these point mutations, small indels (< 4 nt) were overrepresented in the 10% dilution at 30 °C treatment replicates relative to the mutation accumulation experiment, implying that many of these mutations were adaptive during the experiment (table 3). Furthermore, by following the trajectories of the novel variants that arose over time in the core replicates (fig. 5), we observed that numerous variants at 30 °C persisted throughout the experiment until the final time point (n = 26). The mean number of persistent variants above 5% at 30 °C is also nearly four times greater than at 20 °C by the last sequenced time point, and we identified a statistically significant increasing trend of persistent variants over time only at 30 °C (fig. 6, $P < 0.05$, Mann-Kendall test). Such mutations were likely adaptive and beneficial, to different degrees, and allowed the symbionts to improve their fitness and to fine-tune the host interaction at the elevated temperature.

To study whether there are any mutational hotspots in particular genes or pathways, we first focused on qualifying mutations (Deatherage et al. 2017) in the experiment to calculate Dice's Coefficient of Similarity. This allowed us to compare

the similarity of mutations within the same treatment and across different ones, which should provide a good indication as to whether mutations in certain genes or gene pathways are more predominant in one treatment over another. A previous study examined genomic evolution in 30 *E. coli* populations adapting to five different temperature regimes for 2,000 generations (Deatherage et al. 2017). They found that within the same thermal treatment clones shared nearly 17% of their mutated genes, whereas clones under different temperature regimes shared less than 5% of their mutated genes. In our study, we observed a similarity of 5-6% across all replicates at each temperature (10% dilution treatments), and a similarity between the temperature regimes of 2%. Replicates from the same treatment thus shared more mutations amongst each other than replicates from a different treatment, indicating adaptation to the particular experimental conditions. This confirms temperature as an effective selection pressure and further supports that many of the mutations identified are adaptive. The overall lower mutation specificity in our experiment compared to the *E. coli* experiments is consistent with the observed increased genetic diversity due to the structured symbiont populations. In addition, the longer time frame in the previous study may have led to a more pronounced parallelism. In fact, similar calculations were also carried out using the *E. coli* populations from the LTEE (Lenski 2017). After 500 generations the populations shared 15% of affected genes but this rose to 33% at 2,000 and 5,000 generations, which demonstrates that parallelism increased after the first 500 generations.

In total 17 genes were affected by at least two qualifying mutations across the different treatments, and of these, ten genes were restricted to mutations in the 10% dilution treatment at 30 °C (supplementary table S2, Supplementary Material online). Three of these genes encode hypothetical proteins, and three others are predicted to be involved in translation, ribosomal structure and repair. The most prominent of these ten genes was PC_RS06415 (pc1333) because mutations in this gene were identified in five different replicates. PC_RS06415 is similar to known ABC transporter ATP-binding proteins and classified in eggNOG in the category defense mechanisms. While its exact function remains unknown, homologs are interestingly found in related chlamydial organisms and other intracellular bacteria, such as *Legionella* species. Additional genes in this subset include two genes, *ptsI* and *ptsH*, predicted to be involved in the phosphotransferase system PTS, which is incomplete in *P. amoebophila* and likely serves regulatory functions. Two predicted kinases,

phoR and *pknD*, are also affected by qualifying mutations in multiple replicates. The Pho regulon is linked to attenuated virulence and alteration of virulence traits in *E. coli* (Crépin et al. 2011) and to host adaptation in *Mycobacterium tuberculosis* (Chiner-Oms et al. 2019). PknD is well conserved among chlamydiae; its inhibition leads to smaller inclusions and less infective progeny in *C. pneumoniae* (Johnson et al. 2009). Together, qualifying mutations occurred in a number of genes involved in sensing, regulation, and host interaction, suggesting they are involved in the adaptive processes observed.

When we took a broader look at which cellular functions were affected by mutations, genes classified in the categories defense mechanisms and post-translational modification were most pronounced at the elevated temperature (fig. 7 and supplementary table S3, Supplementary Material online). In addition, numerous mutations were identified in eight functional categories related to metabolism, and in the categories translation, ribosomal structure, and biogenesis - including several RNA methyltransferases - as well as replication, recombination and repair. Altogether, the large number of diverse genes and cellular functions affected suggests that adaptation to the elevated temperature was manifold and comprised changes in regulatory pathways, host interaction including metabolite exchange as well as stress response and DNA repair. Yet, although we had observed an overall attenuated infectivity phenotype in the populations evolved at 30 °C (fig. 2 and supplementary fig. S1, Supplementary Material online), we could not identify any mutations in specific genes that are present in a majority of these populations and that could easily explain the reduced infectivity. This might be due to the complex developmental cycle of the symbionts, where changes at various steps of the infection process - during uptake, during establishment of the intracellular niche, during replication, or during secondary differentiation - may fine-tune host interaction and result in a similar outcome, a less infectious phenotype.

Conclusions

Adaptation to elevated temperature is a critical step in the evolution of environmental microbes to pathogens of humans. Here, we investigated how obligate intracellular symbionts of amoebae responded to increased temperature in an evolution experiment spanning 420 to 510 bacterial generations. We found that an increase of incubation temperature from environmental conditions to 30 °C selected for

phenotypes with reduced infectivity in replicate symbiont populations, likely caused by the necessity to balance host-interaction in this challenging environment. This phenotype coincided with numerous mutations in genes involved in diverse cellular functions. Furthermore, the increase in the proportion of nonsynonymous mutations relative to synonymous mutations as well as the overrepresentation of small indels (< 4 nt) indicate that positive selection played a strong role at this elevated temperature.

There is a striking paucity of long-term evolution experiments with strictly intracellular bacteria, perhaps because of several challenges inherent to these systems. This includes experimental limitations, such as the difficulty to establish clonal cultures and the challenges to interpret pool sequencing data, as well as the lack of routine genetic methods for many strictly intracellular bacteria and their protist hosts. In addition, biological features of intracellular microbial symbioses need to be considered, for instance, the coupling of symbiont and host generation times in vertically transmitted symbionts, structured populations as a result of subpopulations being restricted to the intracellular compartment, different levels of selection during the infection process, and the effects of co-evolution with the host. Yet, evolution experiments bear a great potential for improving our knowledge on the evolution of strictly intracellular microbes and microbe-host interaction. As such, the present study uncovered basic features of molecular evolution and temperature adaptation of chlamydial symbionts in their protected intracellular niche, contributing to a better understanding of how these bacteria may have transitioned from symbionts in microbial eukaryotes in the environment to well-adapted and highly successful pathogens of humans.

Acknowledgments

Sequencing was performed at the Vienna Biocenter Core Facility (VBCF) NGS Unit (https://www.viennabiocenter.org/facilities/).

This project has received funding from the European Research Council ERC (EVOCHLAMY, grant no. 281633 to MH) and the Austrian Science Fund FWF in the context of the doc.funds program (DOC 69-B) and the InfectERA project Eugenpath (project number I1628-B22).

MH conceived the study. LS and MH designed the experiments. LS, JS, FW, HN, PH and TK performed the evolution experiment. PH, AR and TK performed the fitness assays. PH, LS and JS performed the genome sequencing experiments. PH, MZ, TK and MH analysed fitness and genome data. TR provided data analysis support. PH and MH wrote the manuscript. All authors edited and approved the final manuscript.

We thank Stephan Köstlbacher for bioinformatics support and help with computational analysis.

Correspondence and material requests should be addressed to MH (matthias.horn@univie.ac.at).

References

Abdelrahman YM, Belland RJ. 2005. The chlamydial developmental cycle. *FEMS Microbiol. Rev.* 29:949–959.

Agashe D, Sane M, Phalnikar K, Diwan GD, Habibullah A, MartinezGomez NC, Sahasrabuddhe V, Polachek W, Wang J, Chubiz LM, et al. 2016. Large-effect beneficial synonymous mutations mediate rapid and parallel adaptation in a bacterium. *Mol. Biol. Evol.* 33:1542–1553. Available from: https://doi.org/10.1093/molbev/msw035.

Anbutsu H, Goto S, Fukatsu T. 2008. High and low temperatures differently affect infection density and vertical transmission of male-killing *Spiroplasma* symbionts in *Drosophila* hosts. *Appl. Environ. Microbiol.* 74:6053–6059.

Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. 2005. Host-bacterial mutualism in the human intestine. *Science* 307:1915–1920.

Balin BJ, Gérard HC, Arking EJ, Appelt DM, Branigan PJ, Abrams JT, Whittum-Hudson JA, Hudson AP. 1998. Identification and localization of *Chlamydia pneumoniae* in the Alzheimer's brain. *Med. Microbiol. Immunol.* 187:23–42.

Barrick JE, Lenski RE. 2009. Genome-wide Mutational Diversity in an Evolving Population of *Escherichia coli*. *Cold Spring Harbor Symp. Quant. Biol.* [Internet] 74:119–129. Available from: http://dx.doi.org/10.1101/sqb.2009.74.018

Barrick JE, Lenski RE. 2013. Genome dynamics during experimental evolution. *Nat. Rev. Genet.* [Internet] 14:827–839. Available from: http://dx.doi.org/10.1038/nrg3564

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461:1243–1247.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.

Borges V, Ferreira R, Nunes A, Sousa-Uva M, Abreu M, Borrego MJ, Gomes JP. 2013. Effect of long-term laboratory propagation on *Chlamydia trachomatis* genome dynamics. *Infect. Genet. Evol.* 17:23–32.

Braendle C, Miura T, Bickel R, Shingleton AW, Kambhampati S, Stern DL. 2003. Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. *PLoS Biol.* 1:E21.

Brockhurst MA, Colegrave N, Rozen DE. 2010. Next-generation sequencing as a tool to study microbial evolution. *Mol. Ecol.* 20:972–980.

Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* 523:208–211.

Buckling A, Craig Maclean R, Brockhurst MA, Colegrave N. 2009. The Beagle in a bottle. *Nature* 457:824–829.

Burch CL, Chao L. 1999. Evolution by small steps and rugged landscapes in the RNA virus phi6. *Genetics* 151:921–927.

Burillo A, Bouza E. 2010. *Chlamydophila pneumoniae*. *Infect. Dis. Clin. North Am.* 24:61–71.

Campbell LA, Kuo C-C. 2004. *Chlamydia pneumoniae* — an infectious risk factor for atherosclerosis? *Nat. Rev. Microbiol.* 2:23–32.

Casson N, Medico N, Bille J, Greub G. 2006. *Parachlamydia acanthamoebae* enters and multiplies within pneumocytes and lung fibroblasts. *Microbes Infect.* 8:1294–1300.

Castelle CJ, Banfield JF. 2018. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* 172:1181–1197.

Chiner-Oms Á, Sánchez-Busó L, Corander J, Gagneux S, Harris SR, Young D, González-Candelas F, Comas I. 2019. Genomic determinants of speciation and spread of the *Mycobacterium tuberculosis* complex. *Sci. Adv.* 5:eaaw3307.

Collingro A, Köstlbacher S, Horn M. 2020. *Chlamydiae* in the Environment. *Trends Microbiol.* [Internet]. Available from: http://dx.doi.org/10.1016/j.tim.2020.05.020

Collingro A, Poppert S, Heinz E, Schmitz-Esser S, Essig A, Schweikert M, Wagner M, Horn M. 2005. Recovery of an environmental *Chlamydia* strain from activated sludge by co-cultivation with *Acanthamoeba* sp. *Microbiology* 151:301–309.

Collingro A, Toenshoff ER, Taylor MW, Fritsche TR, Wagner M, Horn M. 2005. "*Candidatus* Protochlamydia amoebophila", an endosymbiont of *Acanthamoeba* spp. *Int. J. Syst. Evol. Microbiol.* 55:1863–1866.

Conrad TM, Joyce AR, Kenyon Applebee M, Barrett CL, Xie B, Gao Y, Palsson BØ. 2009. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol.* 10:R118.

Crépin S, Chekabab S-M, Le Bihan G, Bertrand N, Dozois CM, Harel J. 2011. The Pho regulon and the pathogenesis of *Escherichia coli*. *Vet. Microbiol.* [Internet] 153:82–88. Available from: http://dx.doi.org/10.1016/j.vetmic.2011.05.043

Cvijović I, Nguyen Ba AN, Desai MM. 2018. Experimental Studies of Evolutionary Dynamics in Microbes. *Trends Genet.* 34:693–703.

Dal Conte I, Mistrangelo M, Cariti C, Chiriotto M, Lucchini A, Vigna M, Morino M, Di Perri G. 2014. Lymphogranuloma venereum: an old, forgotten re-emerging systemic disease. *Panminerva Med.* 56:73–83.

Deatherage DE, Kepner JL, Bennett AF, Lenski RE, Barrick JE. 2017. Specificity of genome evolution in experimental populations of evolved at different temperatures. *Proc. Natl. Acad. Sci. U. S. A.* 114:E1904–E1912.

DeGraves FJ, Gao D, Hehnen H-R, Schlapp T, Kaltenboeck B. 2003. Quantitative

Detection of *Chlamydia psittaci* and *C. pecorum* by High-Sensitivity Real-Time PCR Reveals High Prevalence of Vaginal Infection in Cattle. *J. Clin. Microbiol.* [Internet] 41:1726–1729. Available from: http://dx.doi.org/10.1128/jcm.41.4.1726-1729.2003

Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF, Kassen R. 2012. Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Mol. Ecol.* 21:2058–2077.

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. U. S. A.* 88:7160–7164.

Drake JW. 2009. Avoiding dangerous missense: thermophiles display especially low mutation rates. *PLoS Genet.* 5:e1000520.

Duncan BK, Weiss B. 1982. Specific mutator effects of ung (uracil-DNA glycosylase) mutations in *Escherichia coli*. *J. Bacteriol.* 151:750–755.

Elena SF, Lenski RE. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* 4:457–469.

Elwell C, Mirrashidi K, Engel J. 2016. *Chlamydia* cell biology and pathogenesis. *Nat. Rev. Microbiol.* [Internet] 14:385–400. Available from: http://dx.doi.org/10.1038/nrmicro.2016.30

Fischer A, Rudel T. 2018. Safe haven under constant attack-The *Chlamydia*-containing vacuole. *Cell. Microbiol.* 20:e12940.

Fong SS, Joyce AR, Palsson BØ. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome Res.* 15:1365–1372.

Gerrish PJ, Lenski RE. 1998. The fate of competing beneficial mutations in an asexual population. *Genetica* [Internet]:127–144. Available from: http://dx.doi.org/10.1007/978-94-011-5210-5_12

Gong J, Qing Y, Guo X, Warren A. 2014. "*Candidatus* Sonnebornia yantaiensis", a member of candidate division OD1, as intracellular bacteria of the ciliated protist Paramecium bursaria (Ciliophora, Oligohymenophorea). *Syst. Appl. Microbiol.* 37:35–41.

Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature* 551:45–50.

Gordo I, Campos PRA. 2006. Adaptive evolution in a spatially structured asexual population. *Genetica* [Internet] 127:217–229. Available from: http://dx.doi.org/10.1007/s10709-005-4012-9

Grayston JT, Aldous MB, Easton A, -p. Wang S, -c. Kuo C, Campbell LA, Altman J. 1993. Evidence that *Chlamydia pneumoniae* Causes Pneumonia and Bronchitis. *J. Infect. Dis.* 168:1231–1235.

Greub G, Mege J-L, Raoult D. 2003. *Parachlamydia acanthamoebae* enters and multiplies within human macrophages and induces their apoptosis [corrected].

Infect. Immun. 71:5979–5985.

Habets MGJL, Rozen DE, Hoekstra RF, de Visser JAGM. 2006. The effect of population structure on the adaptive radiation of microbial populations evolving in spatially structured environments. *Ecol. Lett.* 9:1041–1048.

Hadfield J, Harris SR, Seth-Smith HMB, Parmar S, Andersson P, Giffard PM, Schachter J, Moncada J, Ellison L, Vaulet MLG, et al. 2017. Comprehensive global genome dynamics of *Chlamydia trachomatis* show ancient diversification followed by contemporary mixing and recent lineage expansion. *Genome Res.* [Internet] 27:1220–1229. Available from: http://dx.doi.org/10.1101/gr.212647.116

Halligan DL, Keightley PD. 2009. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. *Annu. Rev. Ecol. Evol. Syst.* [Internet] 40:151–172. Available from: http://dx.doi.org/10.1146/annurev.ecolsys.39.110707.173437

Hegreness M, Kishony R. 2007. Analysis of genetic systems using experimental evolution and whole-genome sequencing. *Genome Biol.* 8:201.

Herrera P, Schuster S, Wentrup C, König L, Kempinger T, Na H, Schwarz J, Köstlbacher S, Wascher F, Zojer M, Rattei T, Horn M. 2020. Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont. *Proc. Natl. Acad. Sci. U. S. A.* doi:10.1073/pnas.2005536117.

Horn M. 2008. *Chlamydiae* as symbionts in eukaryotes. *Annu. Rev. Microbiol.* 62:113–131.

Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B, Brandt P, Nyakatura GJ, Droege M, Frishman D, et al. 2004. Illuminating the evolutionary history of Chlamydiae. *Science* 304:728–730.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47:D309–D314.

Hurst GD, Johnson AP, Schulenburg JH, Fuyama Y. 2000. Male-killing *Wolbachia* in *Drosophila*: a temperature-sensitive trait with a threshold bacterial density. *Genetics* 156:699–709.

Hurst GD, Werren JH. 2001. The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.* 2:597–606.

Jia L, Clegg MT, Jiang T. 2003. Excess non-synonymous substitutions suggest that positive selection episodes occurred during the evolution of DNA-binding domains in the *Arabidopsis* R2R3-MYB gene family. *Plant Mol. Biol.* 52:627–642.

Johnson DL, Stone CB, Bulir DC, Coombes BK, Mahony JB. 2009. A novel inhibitor of *Chlamydophila pneumoniae* protein kinase D (PknD) inhibits phosphorylation of CdsD and suppresses bacterial replication. *BMC Microbiol.* 9:218.

Joseph SJ, Didelot X, Rothschild J, de Vries HJC, Morré SA, Read TD, Dean D. 2012. Population genomics of *Chlamydia trachomatis*: insights on drift,

selection, recombination, and population structure. *Mol. Biol. Evol.* 29:3933–3946.

Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012. Experimental evolution. *Trends Ecol. Evol.* 27:547–560.

Kebbi-Beghdadi C, Batista C, Greub G. 2011. Permissivity of fish cell lines to three *Chlamydia*-related bacteria: *Waddlia chondrophila*, *Estrella lausannensis* and *Parachlamydia acanthamoebae*. *FEMS Immunol. Med. Microbiol.* 63:339–345.

König L, Siegl A, Penz T, Haider S, Wentrup C, Polzin J, Mann E, Schmitz-Esser S, Domman D, Horn M. 2017. Biphasic Metabolism and Host Interaction of a Chlamydial Symbiont. *mSystems* [Internet] 2. Available from: http://dx.doi.org/10.1128/mSystems.00202-16

Korona R, Nakatsu CH, Forney LJ, Lenski RE. 1994. Evidence for multiple adaptive peaks from populations of bacteria evolving in a structured habitat. *Proc. Natl. Acad. Sci. U. S. A.* [Internet] 91:9037–9041. Available from: http://dx.doi.org/10.1073/pnas.91.19.9037

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639–1645.

Lebeuf-Taylor E, McCloskey N, Bailey SF, Hinz A, Kassen R. 2019. The distribution of fitness effects among synonymous mutations in a gene under directional selection. *eLife* 8:e45952.

Lee D-H, Palsson BØ. 2010. Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a Nonnative carbon source, L-1,2-propanediol. *Appl. Environ. Microbiol.* 76:4158–4168.

Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 109:E2774–E2783.

Lenski RE. 2017. Convergence and Divergence in a Long-Term Experiment with Bacteria. *Am. Nat.* [Internet] 190:S57–S68. Available from: http://dx.doi.org/10.1086/691209

Lenski RE, Rose MR, Simpson SC, Tadler SC. 1991. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence During 2,000 Generations. *Am. Nat.* 138:1315–1341.

Lenski RE, Wiser MJ, Ribeck N, Blount ZD, Nahum JR, Morris JJ, Zaman L, Turner CB, Wade BD, Maddamsetti R, et al. 2015. Sustained fitness gains and variability in fitness trajectories in the long-term evolution experiment with *Escherichia coli*. *Proc. Biol. Sci.* 282:20152292.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The

Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Lindahl T. 1979. DNA Glycosylases, Endonucleases for Apurinic/Apyrimidinic Sites, and Base Excision-Repair. *Prog. Nucleic Acid Res. Mol. Biol.* [Internet]:135–192. Available from: http://dx.doi.org/10.1016/s0079-6603(08)60800-4

Lind PA, Andersson DI. 2008. Whole-genome mutational biases in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 105:17878–17883.

Maddamsetti R, Lenski RE, Barrick JE. 2015. Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with *Escherichia coli*. *Genetics* 200:619-631. Available from: http://dx.doi.org/10.1101/017020

Maurin M, Bryskier A, Raoult D. 2002. Antibiotic susceptibilities of *Parachlamydia acanthamoeba* in amoebae. *Antimicrob. Agents Chemother.* 46:3065–3067.

Miralles R, Gerrish PJ, Moya A, Elena SF. 1999. Clonal Interference and the Evolution of RNA Viruses. *Science* [Internet] 285:1745–1747. Available from: http://dx.doi.org/10.1126/science.285.5434.1745

Molmeret M, Horn M, Wagner M, Santic M, Abu Kwaik Y. 2005. Amoebae as training grounds for intracellular bacterial pathogens. *Appl. Environ. Microbiol.* 71:20–28.

Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42:165–190.

Nei M, Maruyama T, Chakraborty R. 1975. The Bottleneck Effect and Genetic Variability in Populations. *Evolution* [Internet] 29:1. Available from: http://dx.doi.org/10.2307/2407137

Omsland A, Sixt BS, Horn M, Hackstadt T. 2014. Chlamydial metabolism revisited: interspecies metabolic variability and developmental stage-specific physiologic activities. *FEMS Microbiol. Rev.* 38:779–801.

Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* [Internet] 52:935–949. Available from: http://dx.doi.org/10.1111/j.1558-5646.1998.tb01823.x

Pilalis E, Chatziioannou AA, Grigoroudis AI, Panagiotidis CA, Kolisis FN, Kyriakidis DA. 2011. *Escherichia coli* genome-wide promoter analysis: identification of additional AtoC binding target elements. *BMC Genomics* 12:238.

Polkinghorne A, Borel N, Becker A, Lu ZH, Zimmermann DR, Brugnera E, Pospischil A, Vaughan L. 2009. Molecular evidence for chlamydial infections in the eyes of sheep. *Vet. Microbiol.* 135:142–146.

Poppert S, Essig A, Marre R, Wagner M, Horn M. 2002. Detection and differentiation of Chlamydiae by fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 68:4081–4089.

Rainey PB, Remigi P, Farr AD, Lind PA. 2017. Darwin was right: where now for experimental evolution? *Curr. Opin. Genet. Dev.* 47:102–109.

Rainey PB, Travisano M. 1998. Adaptive radiation in a heterogeneous environment.

Nature [Internet] 394:69–72. Available from: http://dx.doi.org/10.1038/27900

Ray K, Marteyn B, Sansonetti PJ, Tang CM. 2009. Life on the inside: the intracellular lifestyle of cytosolic bacteria. *Nat. Rev. Microbiol.* 7:333–340.

R Development Core Team. 2011. R: a language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.*

Roulis E, Polkinghorne A, Timms P. 2013. *Chlamydia pneumoniae*: modern insights into an ancient pathogen. *Trends Microbiol.* 21:120–128.

Schautteet K, Vanrompay D. 2011. *Chlamydiaceae* infections in pig. *Vet. Res.* [Internet] 42:29. Available from: http://dx.doi.org/10.1186/1297-9716-42-29

Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864.

Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wagner M, Rattei T, Horn M. 2010. The genome of the amoeba symbiont "*Candidatus* Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* 192:1045–1057.

Schulz F, Horn M. 2015. Intranuclear bacteria: inside the cellular control center of eukaryotes. *Trends Cell Biol.* 25:339–346.

Shima K, Kuhlenbäumer G, Rupp J. 2010. *Chlamydia pneumoniae* infection and Alzheimer's disease: a connection to remember? *Med. Microbiol. Immunol.* 199:283–289.

Sixt BS, Hiess B, König L, Horn M. 2012. Lack of effective anti-apoptotic activities restricts growth of *Parachlamydiaceae* in insect cells. *PLoS One* 7:e29565.

Sixt BS, Siegl A, Müller C, Watzka M, Wultsch A, Tziotis D, Montanaro J, Richter A, Schmitt-Kopplin P, Horn M. 2013. Metabolic features of *Protochlamydia amoebophila* elementary bodies--a link between activity and infectivity in *Chlamydiae*. *PLoS Pathog.* 9:e1003553.

Strassmann JE, Shu L. 2017. Ancient bacteria–amoeba relationships and pathogenic animal bacteria. *PLoS Biol.* [Internet] 15:e2002460. Available from: http://dx.doi.org/10.1371/journal.pbio.2002460

Subtil A, Collingro A, Horn M. 2014. Tracing the primordial *Chlamydiae*: extinct parasites of plants? *Trends Plant Sci.* 19:36–43.

Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M. 2012. Drift-barrier hypothesis and mutation-rate evolution. *Proc. Natl. Acad. Sci. U. S. A.* 109:18488–18492.

Taylor-Brown A, Halter T, Polkinghorne A, Horn M. 2020. Thinking Outside the *Chlamydia* Box. *Chlamydia Biology: From Genome to Disease* [Internet]. Available from: http://dx.doi.org/10.21775/9781912530281.18

Taylor-Brown A, Vaughan L, Greub G, Timms P, Polkinghorne A. 2015. Twenty years of research into *Chlamydia*-like organisms: a revolution in our understanding of the biology and pathogenicity of members of the phylum Chlamydiae. *Pathog.*

Dis. 73:1–15.

Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, Wu GC, Wielgoss S, Cruveiller S, Medigue C, et al. 2016. Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature* 536:165-170. Available from: http://dx.doi.org/10.1101/036806

Van den Bergh B, Swings T, Fauvart M, Michiels J. 2018. Experimental Design, Population Dynamics, and Diversity in Microbial Experimental Evolution. *Microbiol. Mol. Biol. Rev.* [Internet] 82. Available from: http://dx.doi.org/10.1128/MMBR.00008-18

Wagner M, Horn M, Daims H. 2003. Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Curr. Opin. Microbiol.* 6:302–309.

Wichman HA, Brown CJ. 2010. Experimental evolution of viruses: Microviridae as a model system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365:2495–2501.

Wiser MJ, Ribeck N, Lenski RE. 2013. Long-term dynamics of adaptation in asexual populations. *Science* 342:1364–1367.

Wright HR, Turner A, Taylor HR. 2008. Trachoma. *Lancet* 371:1945–1954.

Yang Z, Nielsen R. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.* 46:409–418.

Zojer M, Schuster LN, Schulz F, Pfundner A, Horn M, Rattei T. 2017. Variant profiling of evolving prokaryotic populations. *PeerJ* 5:e2997.

Genome dynamics and temperature adaptation during experimental evolution of obligate intracellular bacteria

Paul Herrera, Lisa Schuster, Markus Zojer, Hyunsoo Na, Jasmin Schwarz, Florian Wascher, Thomas Kempinger, Andreas Regner, Thomas Rattei, Matthias Horn

Supplementary information

Table S2: Summary of all genes affected by at least two qualifying mutations across all treatments at the final time point.

NCBI locus tag	Gene name	No. of reps	Which treatment?	RefSeq function	eggNOG functional category
PC_RS06415		5	30	ABC transporter ATP-binding protein	V
PC_RS02320	ptsI	4	20/30	phosphoenolpyruvate-protein phosphotransferase (PTS enzyme I)	G
PC_RS02570		4	30	hypothetical protein	
PC_RS05215	dnaA2	3	20/30	chromosome replication initiator DnaA	L
PC_RS09030	phoR	3	20/30	two-component sensor histidine kinase phoR	T
PC_RS05210	yidC	3	20/30	inner membrane protein translocase component YidC	U
PC_RS05050	vacB	2	20/30	ribonuclease R	K
PC_RS02650		2	20/30	hypothetical protein	R
PC_RS02315	ptsH1	2	20*/30	HPr, histidine-containing phosphocarrier protein of the PTS	G
PC_RS04155		2	30	hypothetical protein	S
PC_RS05410		2	30	HlyC/CorC family transporter	P
PC_RS07875	pknD	2	30	serine/threonine protein kinase	T
PC_RS02040	rpsC	2	30	30S ribosomal protein S3	J
PC_RS04415	kefC	2	30	glutathione-regulated potassium-efflux system protein	
PC_RS01330	radA	2	30	DNA repair protein RadA	L
PC_RS01205		2	30	RNA methyltransferase	J
PC_RS07790		2	30	hypothetical protein	

*All mutations here were identified in the 10% dilution treatments at either temperature, except for one mutation in PC_RS02315 identified in the 0.1% dilution treatment at 20 °C.

Table S3: Summary of all genes with known functional categories affected by mutations identified in the final sequenced time point across both temperatures in the 10% dilution treatment.

NCBI locus tag	Gene name	Which treatment?	RefSeq function	eggNOG functional category
PC_RS01145	dnaQ1	20	DNA polymerase III subunit epsilon	L
PC_RS02805	NA	20	hypothetical protein	P
PC_RS03460	prpC	20	serine/threonine phosphatase	T
PC_RS03510	NA	20	hypothetical protein	O
PC_RS05990	miaA	20	tRNA (adenosine(37)-N6)-dimethylallyltransferase	J
PC_RS07690	rpmJ	20	50S ribosomal protein L36	J
PC_RS00115	ahpC	30	thioredoxin peroxidase	O
PC_RS00150	groEl1	30	molecular chaperone	O
PC_RS00480	NA	30	putative inorganic phosphate transporter	P
PC_RS01025	rpsL	30	30S ribosomal protein S12	J
PC_RS01205	NA	30	putative RNA methyltransferase	J
PC_RS01210	recJ	30	single-stranded-DNA-specific exonuclease	L
PC_RS01330	radA	30	DNA repair protein RadA	O
PC_RS01660	rlmN1	30	50S rRNA methyltransferase	J
PC_RS02040	rpsC	30	30S ribosomal protein S3	J
PC_RS02275	dnaJ	30	molecular chaperone	O
PC_RS02310	hprK	30	HPr kinase/phosphorylase	T
PC_RS02315	ptsH1	30	phosphocarrier protein HPr	G
PC_RS02380	NA	30	hypothetical protein	L
PC_RS02640	bepE	30	multidrug efflux RND transporter permease subunit	P
PC_RS02735	nuoN	30	NADH-quinone oxidoreductase subunit N	C
PC_RS02845	infA	30	translation initiation factor IF-1	J
PC_RS02900	rpoC	30	DNA-directed RNA polymerase subunit beta'	K
PC_RS02975	xseA	30	exodeoxyribonuclease VII large subunit	L
PC_RS02980	dagA1	30	amino acid transporter	E
PC_RS03015	ypwA	30	carboxypeptidase M32	E
PC_RS03035	NA	30	hypothetical protein	V
PC_RS03085	rpsO	30	30S ribosomal protein S15	J
PC_RS03530	NA	30	hypothetical protein	O
PC_RS03665	rpsA	30	30S ribosomal protein S1	J
PC_RS03670	lysC	30	aspartate kinase	E
PC_RS04345	htpX	30	peptidase	O
PC_RS04415	kefC2	30	glutathione-regulated potassium-efflux system protein	P
PC_RS05080	NA	30	hypothetical protein	P
PC_RS05320	asd	30	aspartate-semialdehyde dehydrogenase	E
PC_RS05410	NA	30	hypothetical protein	P
PC_RS05575	folB	30	dGTP pyrophosphohydrolase/dihydroneopterin aldolase	F
PC_RS05975	glnA	30	glutamine synthetase	E

NCBI locus tag	Gene name	Which treatment?	RefSeq function	eggNOG functional category
PC_RS06050	NA	30	eukaryotic stearoyl-CoA 9-desaturase	I
PC_RS06290	pcnB2	30	polynucleotide adenylyltransferase	J
PC_RS06300	dpm1	30	glycosyltransferase	M
PC_RS06415	NA	30	ABC transporter ATP-binding protein	V
PC_RS06485	RNR1	30	ribonucleoside-diphosphate reductase subunit alpha	F
PC_RS06590	NA	30	helicase	L
PC_RS06865	NA	30	hypothetical protein	L
PC_RS06880	traC	30	type-IV secretion system protein	U
PC_RS07800	cydA	30	cytochrome ubiquinol oxidase subunit I	C
PC_RS07875	pknD	30	serine/threonine protein kinase	T
PC_RS07950	menC	30	o-succinylbenzoate synthase	H
PC_RS07975	atpD1	30	ATP synthase subunit beta	C
PC_RS08170	recF	30	DNA replication and repair protein	L
PC_RS08310	NA	30	YggS family pyridoxal phosphate enzyme	F
PC_RS08400	NA	30	hypothetical protein	Q
PC_RS08435	glgB	30	glycogen-branching enzyme	G
PC_RS08925	NA	30	MFS transporter	G
PC_RS09395	NA	30	hypothetical protein	U
PC_RS09505	NA	30	SAM-dependent methyltransferase	Q
PC_RS09585	rumA	30	RNA methyltransferase	J
PC_RS09655	ung	30	RNA methyltransferase	L
PC_RS09795	plsC	30	1-acyl-sn-glycerol-3-phosphate acyltransferase	I
PC_RS02320	ptsI	20/30	phosphoenolpyruvate-protein phosphotransferase	G
PC_RS05050	rnr1	20/30	ribonuclease R	K
PC_RS05210	yidC	20/30	membrane protein insertase	U
PC_RS05215	dnaA2	20/30	Chromosomal replication initiator protein	L
PC_RS08765	NA	20/30	hypothetical protein	T
PC_RS09030	phoR	20/30	histidine kinase	T

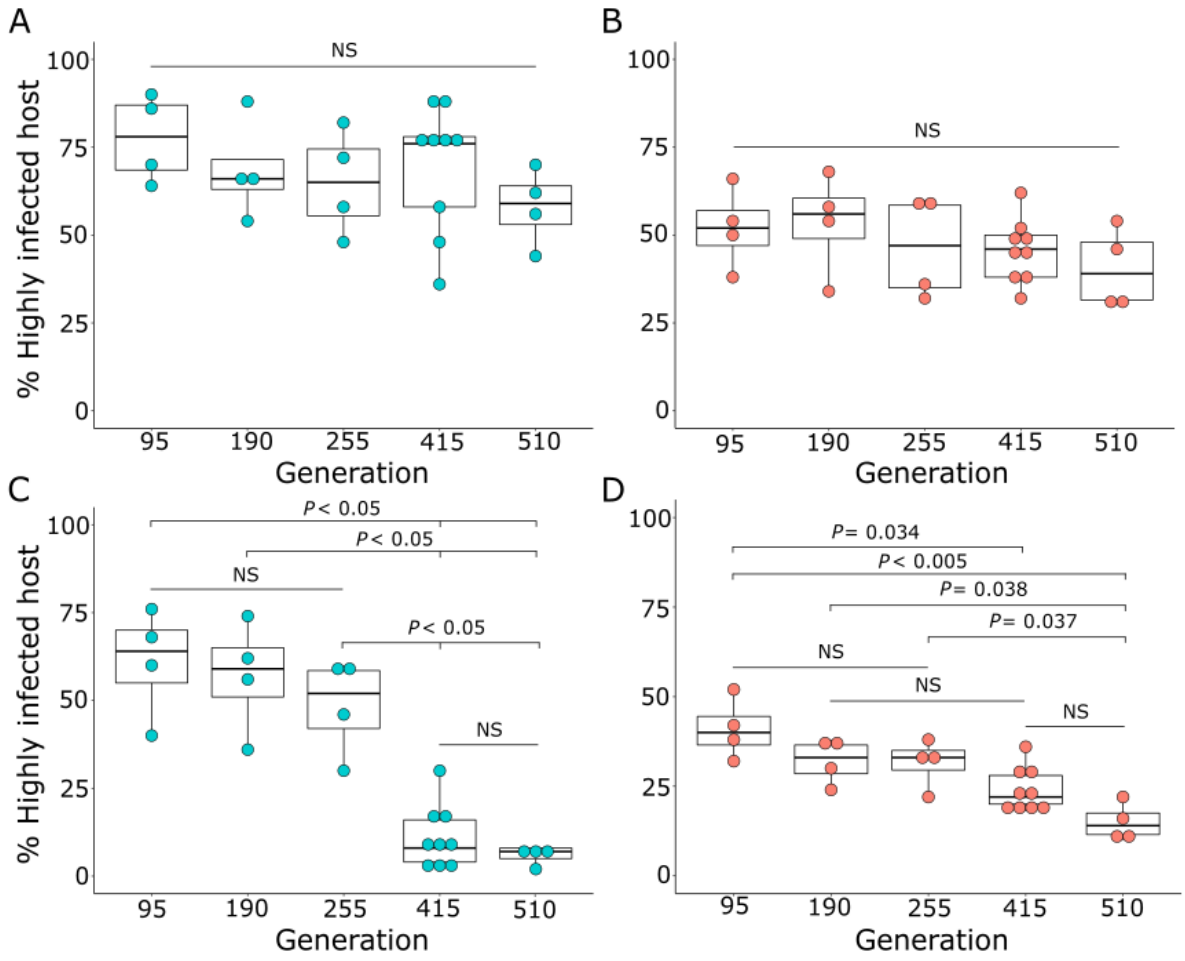

FIG. S1. Time series analysis of symbionts isolated from the 10% dilution populations at both temperatures. Naive ancestral amoebae (*n* = 4 for all generations except at 415 generations, *n* = 9) at a concentration of 10^5 cells per ml were infected with symbionts isolated from evolved populations (10% dilution) at different time points in the evolution experiment with a multiplicity of infection (MOI) of 30, at an assay temperature of 20 °C (blue dots) and at 30 °C (red dots). The percent of highly infected amoebae was assessed at 96 hpi using fluorescence *in situ* hybridization. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. **(*A* and *B*)** Infectivity assays performed using EBs isolated from the 20 °C treatment. The infectivity across the generations was compared using a Kruskal-Wallis test. **(*C* and *D*)** Infectivity assays performed using EBs isolated from the 30 °C treatment. The infectivity across the generations was compared using a Kruskal-Wallis test (C $\chi^2 = 18.632$, $p < 0.005$; D $\chi^2 = 15.721$, $p < 0.005$) followed by Dunn's multiple comparison post-hoc test.

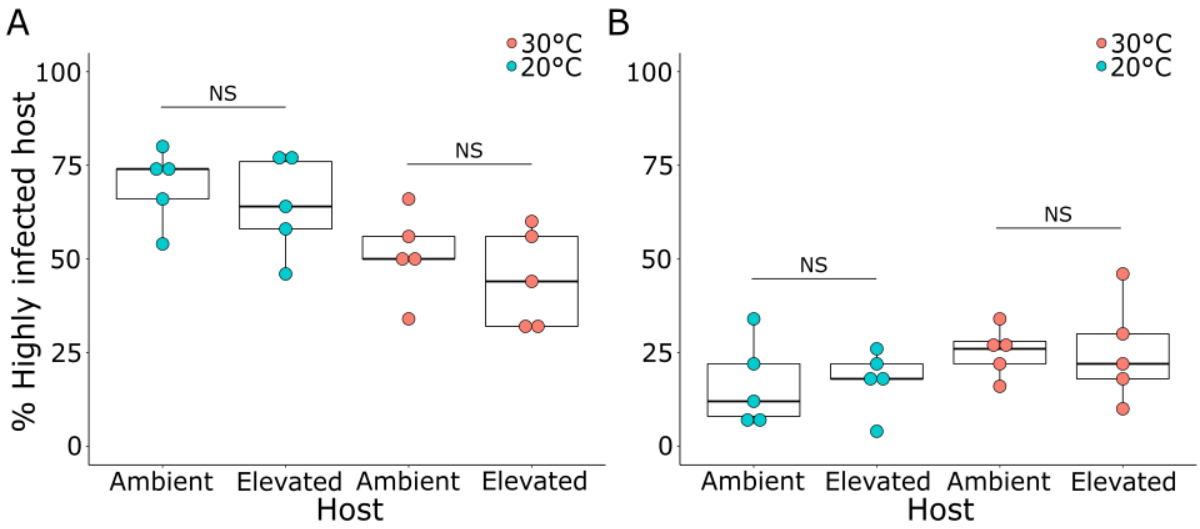

FIG. S2. Infectivity of symbionts did not change when infecting hosts maintained at ambient (20 °C) or elevated (30 °C) temperature. Amoebae maintained at 20 °C or at 30 °C (*n* = 5 for each) at a concentration of 10^5 cells per ml were infected with symbionts isolated from evolved populations (10% dilution) after 415 generations with an MOI of 30, at an assay temperature of 20 °C (blue dots) and at 30 °C (red dots). The percent of highly infected amoebae was assessed at 96 hpi using fluorescence *in situ* hybridization. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. **(*A*)** Infectivity assays performed using EBs isolated from the 20 °C treatment. The infectivity between the two hosts at the same temperature was compared using a two-sided Wilcoxon–Mann–Whitney test. **(*B*)** Infectivity assays performed using EBs isolated from the 30 °C treatment. The infectivity between the two hosts at the same temperature was compared using a two-sided Wilcoxon–Mann–Whitney test.

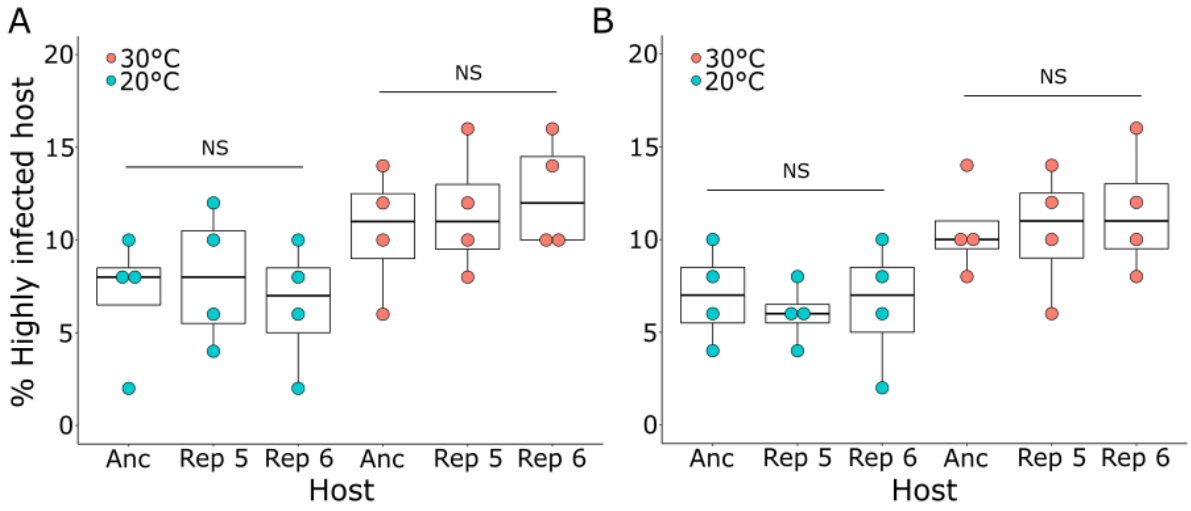

FIG. S3. Infectivity of symbionts did not change when re-infecting their co-evolved hosts. Naive ancestral amoebae, cured amoebae from the same population and cured amoebae from a different population (n = 4 for each) at a concentration of 10^5 cells per ml were infected with symbionts isolated from two 30 °C evolved populations (10% dilution replicates 5 and 6) after 510 generations with an MOI of 30, at an assay temperature of 20 °C (blue dots) and at 30 °C (red dots). The percent of highly infected amoebae was assessed at 96 hpi using fluorescence *in situ* hybridization. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. **(*A*)** Infectivity assays performed using EBs isolated from replicate 5. The infectivity among the host groups was compared using a Kruskal-Wallis test for each temperature. **(*B*)** Infectivity assays performed using EBs isolated from replicate 6. The infectivity among the host groups was compared using a Kruskal-Wallis test for each temperature.

Chapter V

Survival at an elevated temperature strongly reduces gene expression dynamics in a chlamydial symbiont

Title:

Survival at an elevated temperature strongly reduces gene expression dynamics in a chlamydial symbiont

Authors:

Paul Herrera[1], Cecilia Wentrup[1], Lena König[1], Stephan Köstlbacher[1], Thomas Rattei[1], Matthias Horn[1]

Author affiliations:

[1] Centre for Microbiology and Environmental Systems Science, University of Vienna

Corresponding author:

Univ.-Prof. Dr. Matthias Horn

University of Vienna

Department of Microbiology and Ecosystem Science

Althanstr. 14

A-1090 Wien

Austria

phone: +43 1 4277 91208

email: matthias.horn@univie.ac.at

Abstract

Bacteria must constantly adapt to any fluctuations in their surrounding environment to survive and thrive. One important adaptation, especially for those microbes that invade warm-blooded animals, is temperature. This can oscillate dramatically between the environment and the host. Bacteria often increase the expression of temperature-regulated genes to adapt to these changes. We have performed a set of RNA sequencing experiments on selected amoeba-symbiont populations at an elevated temperature (30 °C). One of these populations, which exhibited a reduced infectivity, was isolated from an evolution experiment that was itself maintained at that same temperature. These symbionts live ubiquitously in free-living amoebae and share a common ancestor with major human pathogens. We wanted to obtain initial insights into how these bacteria alter their gene expression dynamics in response to an increase in temperature. We observed shallow expression dynamics as a result of incubation at an elevated temperature, but even lower expression levels when symbionts had been maintained at an elevated temperature for a long period of time. A number of pathways crucial for effective infection and development were not upregulated, in particular those involved in virulence as well as carbon and energy metabolism.

Introduction

Understanding how bacteria adapt to stressful situations and overcome challenges is crucial in today's world, both due to current environmental issues such as climate change, but also due to the never ending arms race between our immune system and pathogens. Temperature adaptation is of particular interest because it is such a crucial step for bacteria to make the leap from environmental reservoirs to vertebrates, and eventually become major human pathogens. Such adaptations tend to consist of increases in the expression of temperature-regulated genes. Frequently this requires global gene regulation involving a two-component signal transduction system made up of a sensor-kinase and a response regulator (Konkel and Tilly 2000). An advantage of this is that it allows unlinked genes to be simultaneously expressed. Global networks for bacterial gene regulation are not only specific to temperature adaptation. They also govern processes such as spore formation (Fawcett et al. 2000) as well as cell cycle developmental changes (Laub et al. 2000).

Previous work carried out by our lab (Herrera, Schuster, Zojer, et al. 2020) sought to gain insight into how obligate intracellular microorganisms adapt to elevated temperature by setting up an evolution experiment. More specifically, we established replicate *Protochlamydia amoebophila* populations thriving in their *Acanthamoeba* hosts under two temperature regimes (20 °C, 30 °C) for 420 to 510 generations (38 months). *Protochlamydia* is just one example of the so-called environmental chlamydiae (Horn 2008) that thrive in the environment, with the majority of them considered to be symbionts of free-living amoebae (Maurin et al. 2002; Collingro, Poppert, et al. 2005; Sixt et al. 2012). These share a common ancestor with major bacterial pathogens of humans, and this is particularly intriguing since temperature adaptation was a critical step in the transition from protist-dwellers to infecting human cells. We were able to track the evolutionary dynamics of this chlamydia symbiont within its protected intracellular niche. By carrying out pool sequencing, coupled with variant detection, we have observed various genomic adaptations. Briefly, we identified evidence for positive selection for both temperatures, but this was more pronounced at the elevated temperature. Symbionts from 30 °C populations had an increase in the proportion of nonsynonymous mutations relative to synonymous mutations, and an overrepresentation of small indels. Genes that were affected by mutations strongly differed with respect to their cellular functions between the two temperature treatments. By comparing infectivity of ancestral and evolved populations, those at 30 °C exhibited attenuated infectivity, possibly as a trade-off to reduce host cell burden.

Although we observed these interesting genomic and phenotypic adaptations, we were not satisfied and wanted to improve our understanding of the molecular changes that allowed the symbionts to adapt to the elevated temperature. We decided that, in order to add another piece to this puzzle, it would be useful to observe gene expression dynamics of these evolved populations. One way for the symbionts to survive the increased temperature could be to modify levels of gene transcription in particular pathways. We therefore performed two RNA sequencing (RNA-Seq) experiments on symbionts infecting naive amoebae at the elevated temperature of 30 °C, isolating RNA at selected critical time points that mark crucial developmental events during infection: 2 hours post infection (hpi) representing the start of infection and elementary body (EB) to reticulate body (RB) transition; 48 hpi as the peak of replication and RB activity; and released EBs accumulated at 7 days

post infection, representing the new generation of symbionts ready to infect new host cells. We performed the RNA-Seq experiments on two populations of chlamydiae. One of these populations was the ancestral chlamydia population that was the founder of the evolution experiment. This had been accustomed to growth at 20 °C, and an RNA-Seq experiment had already been carried out on it (König et al. 2017). The second chlamydia population had to be chosen from the replicates that were adapted to 30 °C during the evolution experiment for 510 generations. The replicate chosen was that one in particular that possessed mutations in genes that were thought to potentially be important for temperature adaptation, which is discussed in greater detail later.

These RNA-Seq experiments thus allowed us to explore two questions: (a) What effect does the incubation temperature have on gene expression dynamics?; and (b) Has prolonged evolution at 30 °C brought about adaptation on a transcriptional level? We attempted to answer these two questions by making two comparisons. First, by comparing the RNA-Seq data obtained by König et al. in 2017 on the ancestral symbionts at 20 °C to the data obtained by us at 30 °C using the same symbionts, we were able to identify whether simply carrying out the experiment at an elevated temperature altered gene expression dynamics. Second, by comparing the two RNA-Seq data sets obtained by us, we improved our understanding of the effect that prolonged cultivation at an elevated temperature had on gene expression patterns. We show that elevated temperature strongly reduced gene expression dynamics, in particular with respect to long-term survival. This is especially true at the early stage of infection for essential pathways involved in virulence and energy metabolism.

Results and Discussion

Overall differential gene expression levels

In order to obtain a preliminary picture, we first looked at the numbers of differentially expressed (DE) genes (supplementary table S1, Supplementary Material online) between two consecutive time points for our three separate data sets (Fig. 1). When carrying out the experiment at the lower temperature, the number of DE genes for each two-way comparison was very similar, with a mean of 440 genes. This would imply that many genes and pathways are important at specific time points throughout the developmental cycle, and are therefore differentially expressed. However, we did

not see the same picture when the experiment was carried out at the higher experiment temperature. In particular, there was very low (20 genes for ancestral symbionts) to no (for evolved symbionts) differential gene expression between 48 hpi and released EBs (Fig. 1). When we look at the other two comparisons (i. e. 2 hpi and 48 hpi; released EBs and 2 hpi) and compare them to the experiment at 20 °C, there are around 60% the number of DE genes for the ancestor and only 15% the number of DE genes for the evolved population. These lower values already provide an early indication that gene expression dynamics are actually strongly reduced both as a result of incubation temperature as well as due to prolonged cultivation at an elevated temperature.

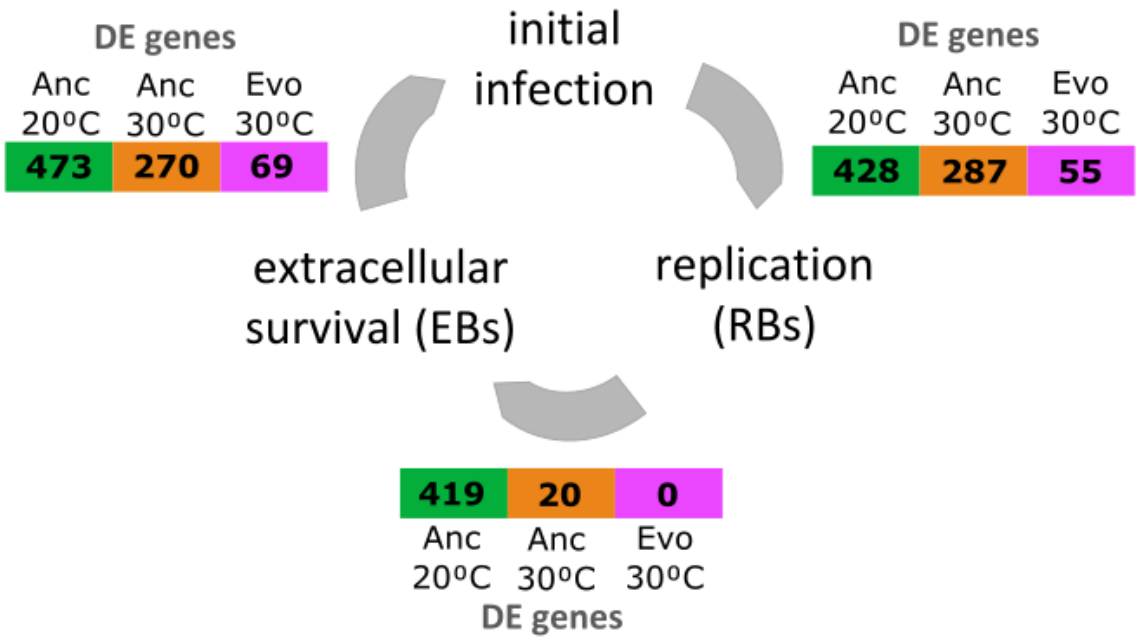

FIG. 1. An overview of the total number of differentially expressed (DE) genes for the three experiments between two consecutive time points that marked crucial developmental events during infection. Genes were considered differentially expressed if their expression changed twofold with a false-discovery rate (FDR) smaller or equal 0.05.

We also compared the gene expression profiles of the ancestor and evolved symbionts at the elevated temperature (Fig. 2 and supplementary table S2, Supplementary Material online). We observed pronounced temporal signatures for the evolved symbionts as has been previously reported by König et al. (2017) and in another chlamydial symbiont (Herrera, Schuster, Wentrup, et al. 2020), whereas for the ancestral symbionts there was a very similar transcriptional signature between the later two time points. The most striking difference in gene regulation between the

two chlamydial symbionts is evident at 2 hpi. This time point is also highly separated from the other two time points for both the ancestral and the evolved symbionts. One could argue that at 2 hpi, gene expression may still be mostly dependent on the EBs that were introduced into the experiment. So the ancestral symbionts maintained at 20 °C presented a particular gene expression pattern, whereas the evolved symbionts maintained at 30 °C presented a different pattern. However, at 48 hpi and in the released EBs, the experiment temperature would definitely have a much stronger effect, and in fact the other two time points present a more similar signature, especially for the released EBs. In order to obtain a clearer idea of what may be occurring at the functional level, we next looked more closely at specific pathways crucial for chlamydial development.

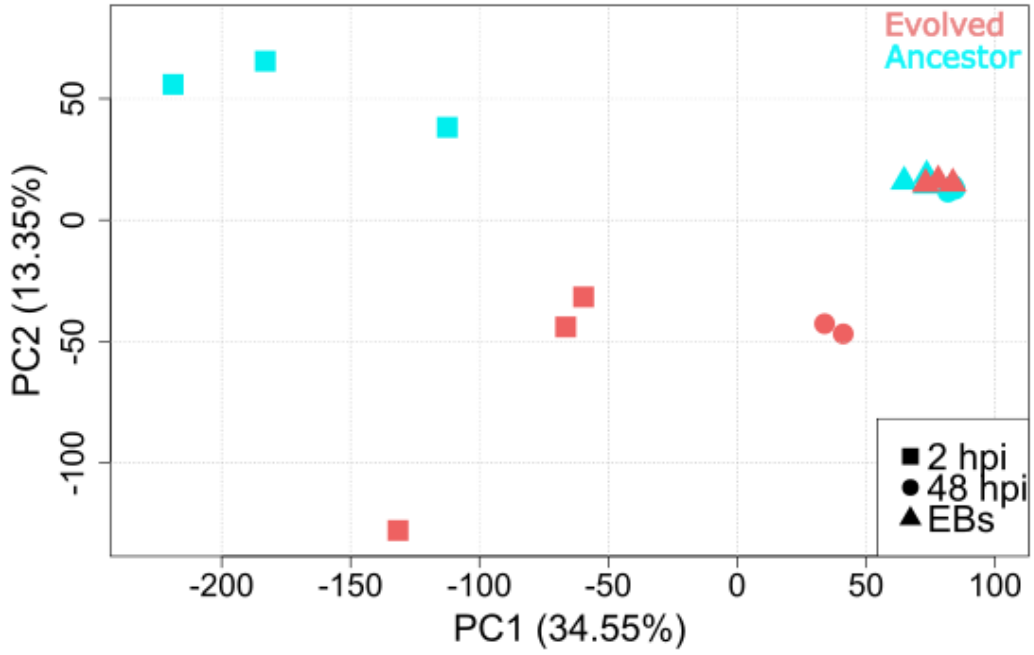

FIG. 2. Principal component (PC) analysis plot for RNA-Seq gene expression data at three time points of symbionts from the ancestor and evolved populations at 30 °C that were collected from amoeba host cells (2 hpi and 48 hpi) or from the medium (released EBs) (*n* = 3, except for evolved 48 hpi where *n* = 2).

Chlamydial type III secretion system

The type III secretion system (T3SS) is a major virulence factor that is evolutionarily well conserved and was already present in the last common ancestor of all known chlamydiae (Betts-Hampikian and Fields 2010; Subtil et al. 2014). König et al. (2017) observed that most genes involved in T3SS were expressed later in the infection during maturation of EBs as well as early during infection. These genes include

structural components, chaperones as well as effectors. In the ancestor at 30 °C we observe a very similar DE gene pattern early in the infection cycle, with many T3SS genes upregulated at 2 hpi. However, there is a stark contrast in the DE genes for the released EBs, with the signature upregulation of T3SS genes being predominantly absent. For the evolved chlamydiae, the gene expression dynamics differ not only for the late expressed genes of the T3SS but also, to a certain extent, at 2 hpi. This is most evident for locus 5 (Fig. 3), the largest locus of the T3SS. Practically all the genes crucial for the T3SS (structural components, chaperones and effectors) were not upregulated at 2 hpi for the evolved symbionts, and this suggests that the initial infection is not being carried out as effectively as in the ancestral strain.

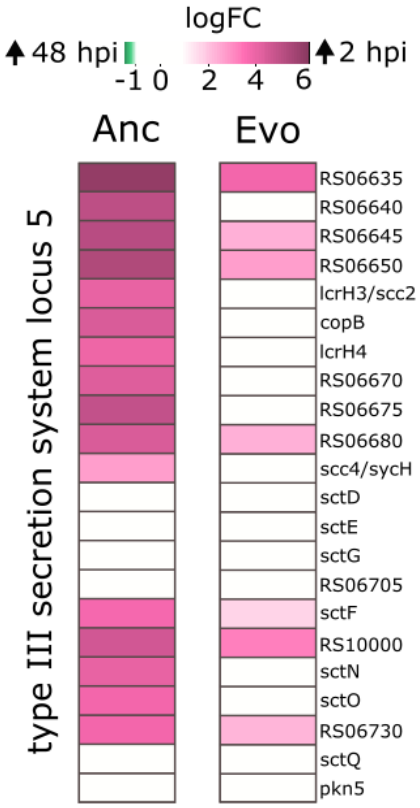

FIG. 3. Heat map depicts the T3SS locus 5 with marked differences between the ancestor and evolved populations at 30 °C with respect to $\log_2$ fold changes (logFC) of genes that were differentially expressed between 2 hpi and 48 hpi (false discovery rate of 0.05). A logFC below 1 (no differential expression) is shown in white and significant upregulation at 2 hpi is shown in pink. The stronger the pink color, the stronger the gene expression change between the two time points. Genes in locus 5 encode chaperones, effectors and structural components of the T3SS.

Carbon and energy metabolism

Chlamydiae are able to synthesize ATP through oxidative phosphorylation. König et al. (2017) observed that genes involved in carbon and energy metabolism are mainly induced late during EB maturation - this included respiratory chain genes as well as genes involved in glycolysis, the tricarboxylic acid cycle (TCA) and the pentose phosphate pathway. These genes are also upregulated, to a lesser extent, at 2 hpi since this initial infection phase is also predominantly composed of EBs. It is clear that from the start of the infection, the EBs in the 30 °C experiments are not functioning in the way that has previously been reported, and this is most apparent when looking at these genes involved in carbon and energy metabolism. In both RNA-Seq datasets, there is no upregulation of these genes at 2 hpi (supplementary table S2, Supplementary Material online). Even more striking is the fact that practically the same situation was observed late during development. None of these genes were upregulated late for the evolved symbionts, whereas only a handful did so for the ancestor, mainly a few respiratory chain genes. EBs, the extracellular stage, are required for survival and infection of new host cells. The reduced expression of these previously mentioned genes, all involved in the utilization of carbon substrates and energy conservation, implies that the EBs are not primed for extracellular survival or new host infection, which are crucial stages for these bacterial symbionts.

Mid-cycle activity

Apart from being able to synthesize ATP via oxidative phosphorylation, chlamydiae are known to be energy parasites, importing host-derived ATP in exchange for ADP via nucleotide transport proteins (Tjaden et al. 1999; Haferkamp et al. 2006). In fact, this is the most important source of energy mid-cycle during intense RB replication. The ATP/ADP translocase (ntt1) is constitutively expressed at a very high level, with its transcriptional profile peaking at mid-cycle, for all three experiments at both temperatures. This demonstrates how crucial the import of host-derived ATP is for chlamydial development. A similar pattern has also been shown for *Chlamydia pneumoniae* and *C. trachomatis* (Belland et al. 2003; Mäurer et al. 2007). Cellular processes such as translation and metabolic activity were observed to be at their most prominent at mid-cycle during the peak of RB activity when performing the RNA-Seq experiment at the lower temperature (König et al. 2017). In fact, there was

a significant functional enrichment among DE genes in categories involved in translation, as well as amino acid and lipid transport and metabolism, at this time point. Furthermore, numerous genes responsible for cell division and peptidoglycan synthesis were upregulated at 48 hpi. The RNA-Seq experiments that we carried out at the elevated temperature did not follow this same metabolic pattern. No functional categories involved in translation and metabolic processes were significantly enriched, which is in stark contrast to the findings from the 20 °C experiment. Nevertheless, there was an upregulation of genes involved in amino acid transport and metabolism for the ancestral symbionts at 48 hpi, and to a lesser extent also for the evolved symbionts (Fig. 4*A* and supplementary table S2, Supplementary Material online). A similar pattern was observed for genes essential for cell division and peptidoglycan synthesis (Fig. 4*B* and supplementary table S2, Supplementary Material online). Overall, these results imply that there is an attenuation of RB activity, with the increased temperature appearing to diminish chlamydial development.

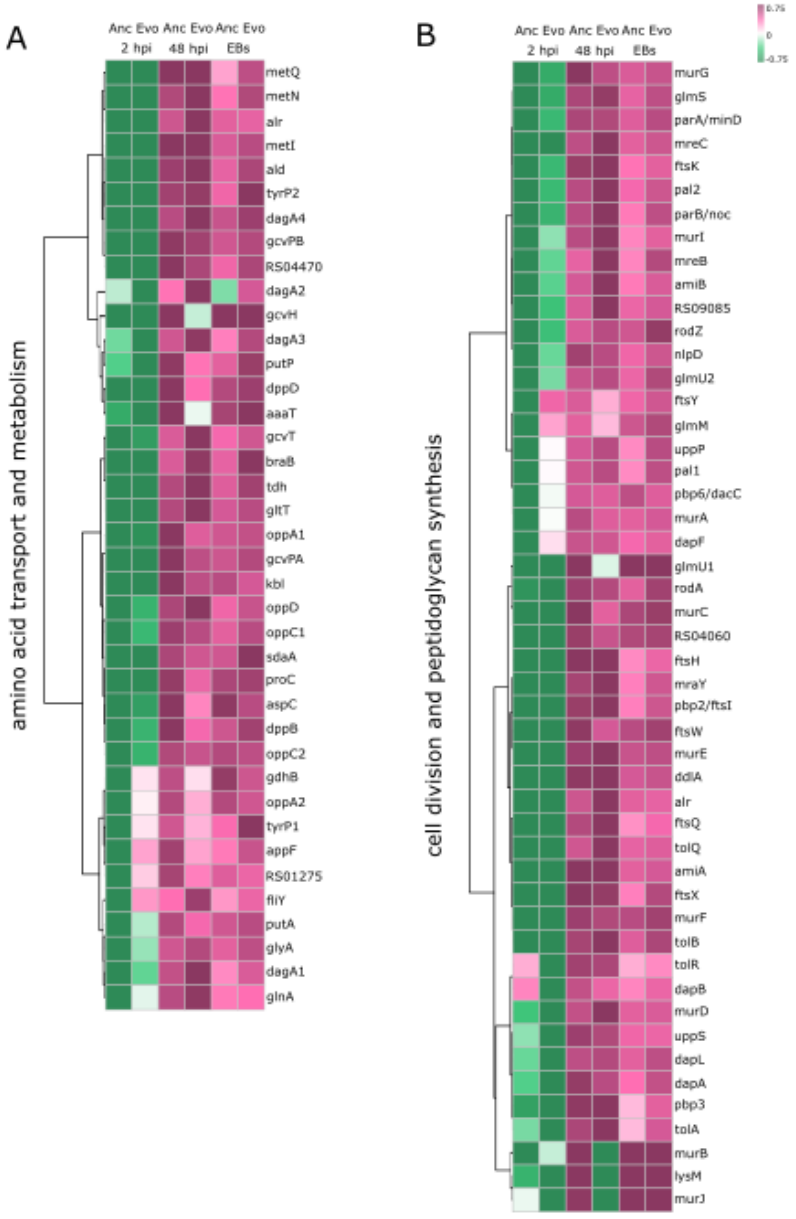

FIG. 4. Heat maps of RNA-Seq data from mean centered gene expression (TPM) of symbiont genes in the ancestor and evolved populations at 30 °C at all three time points during infection. All known loci involved or putatively involved in (A) amino acid transport and metabolism, and (B) cell division and peptidoglycan synthesis are depicted. Highly expressed genes are shown in pink, whereas low gene expression is shown in green. The stronger the color, the stronger the gene expression value.

Mutations identified in the evolved population

The particular population of evolved symbionts used for the RNA-Seq experiment at the elevated temperature was chosen due to the mutations that had accumulated in this population throughout the duration of the evolution experiment (supplementary table S3, Supplementary Material online). The variants that had seemed most interesting and that could have had an adaptive role to survive at the elevated

temperature were those that were located in genes that functioned as kinases. A kinase is an enzyme that catalyzes the transfer of phosphate groups from high-energy, phosphate-donating molecules, such as ATP, to specific substrates in a process known as phosphorylation. The histidine kinase phoR is associated with a reduction in virulence and modification of virulence traits in *E. coli* (Crépin et al. 2011) and to host adaptation in *Mycobacterium tuberculosis* (Chiner-Oms et al. 2019). The serine/threonine-protein kinase pknD is well conserved among chlamydiae; its inhibition results in smaller inclusions as well as less infective progeny in *C. pneumoniae* (Johnson et al. 2009). Since they were both present as nonsynonymous variants in the respective genes, they could have had an adaptive effect to allow the symbionts to improve their chances of survival at the higher temperature. However, when we identified all the DE genes in our experiment, we did not see these kinases in either of the comparisons. This implies that if the variants were having an effect, it does not appear to be related to the gene expression dynamics of these particular genes. In spite of this, phoR may be involved in the regulation of numerous other genes. If this is the case, even though phoR was not differentially expressed, minor differences in its expression levels may have led to pronounced changes in the expression of other genes that it regulates. From all the identified genes that acquired mutations throughout the experiment of this particular replicate, only two of them were observed as being differentially expressed in the experiment: the 23S rRNA gene (RS04995) and the potential MFS transporter (RS08925) were both upregulated at 2 hpi.

Although we cannot explain the observed differences in gene expression by the variants alone, this does not mean that they did not have any impact at all. An important factor to keep in mind when discussing regulation of gene expression in prokaryotes is the presence of regulatory proteins. These DNA-binding proteins recognize specific motifs present in promoter regions and can either repress or activate transcription via interactions with RNA polymerase. These transcription factors may be specific to one gene or to a number of genes. In our study, symbionts evolved at 30 °C may have experienced an increase in repressor proteins, silencing groups of genes that ultimately resulted in these shallower expression dynamics. Most regulatory genes in *Protochlamydia* are, in fact, yet to be identified. Apart from this, a number of repressors have a fine-tuning system termed attenuation. This

makes use of the structure of mRNA to terminate both transcription and translation based on the concentration of an operon's end-product enzymes.

Infectivity assays

We have already carried out numerous infectivity assays on the ancestral symbionts as well as the symbionts isolated from the evolution experiment at 20 °C and at 30 °C (Herrera, Schuster, Zojer, et al. 2020). Here we show those infectivity assays that are pertinent to the three RNA-Seq experiments discussed in this manuscript, i.e. the assays performed on the ancestral symbionts and on the symbionts isolated from the population evolved for 510 generations at 30 °C that was used for the RNA-Seq experiment (Fig. 5). The percent of amoebae highly infected with ancestral symbionts at 96 hpi (i.e. after the development cycle of the symbiont is complete) was very high at the lower temperature of 20 °C, with a median of 80%. When the experiment was carried out at the higher temperature of 30 °C, the percentage of highly infected amoebae was significantly lower, with a median of 50%. However, each of these values was significantly higher than the respective values of the evolved symbionts, both with medians lower than 15% highly infected amoebae at 96 hpi.

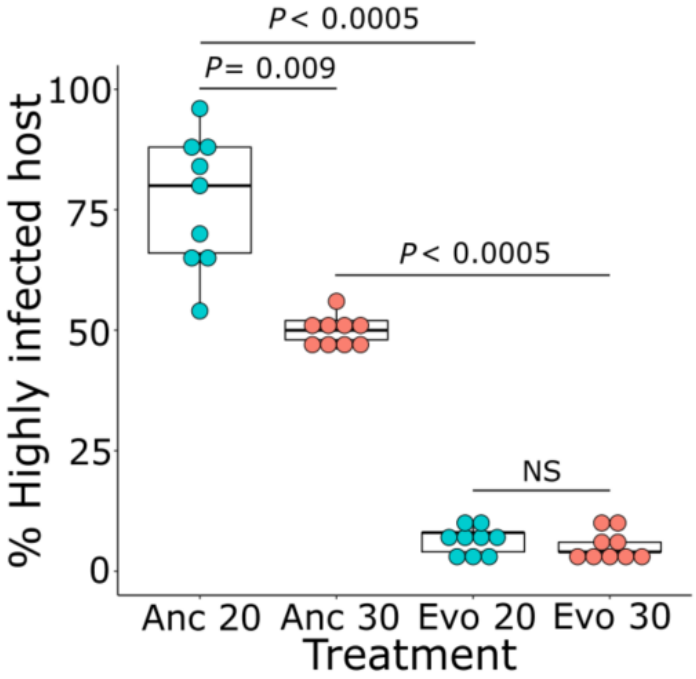

FIG. 5. Summary of infectivity assays at 20 °C and 30 °C. Naive amoebae (*n* = 9 for ancestor, *n* = 5 for evolved population) at a concentration of 10^5 cells per ml were infected with symbionts isolated from the ancestor and the 30 °C evolved population used for the

RNA-Seq experiment with a multiplicity of infection (MOI) of 30. Infectivity assays were carried out at 20 °C (blue dots) and at 30 °C (red dots). The percent of highly infected amoebae was assessed at 96 hpi using fluorescence in situ hybridization. In the box plots, the interquartile range (IQR) between the first and the third quartiles is indicated by the box, while vertical lines extend to a distance of 1.5 x IQR from the first or third quartile. The horizontal line within the box represents the median. The infectivity between the two treatments at the same temperature was compared using a two-sided Wilcoxon–Mann–Whitney test. The infectivity between the same treatment at different temperatures was compared using a Wilcoxon signed-rank test.

We attempted to link these differences in infectivity to the different gene expression dynamics obtained from the RNA-Seq experiments. At 20 °C the ancestral symbionts were the most infectious, and this correlated very well with the transcriptomic data. EBs, the infectious particles that have to enter the host cell, need to be properly primed for an infection, and this was observed for the ancestral symbionts at 20 °C (König et al. 2017). Genes involved in the T3SS and genes important in carbon and energy metabolism were heavily upregulated in the EBs, allowing for the initial infection to be successful. Apart from this, genes essential for RB replication were highly upregulated at mid-cycle, continuing the infection process as required. As a result of this, a very high level of infection was observed by the end of the chlamydial developmental cycle at 96 hpi.

The ancestral symbionts at 30 °C had a lower infectivity (Fig. 5), and this suggests that the gene expression dynamics may differ to the ones observed when the experiment was carried out at the lower temperature. At 2 hpi there was a very similar DE pattern with respect to most T3SS virulence genes, suggesting that the initial infection is not too different from what we would expect. But, apart from this, we have described how different sets of genes crucial for numerous pathways that were upregulated at 20 °C were not upregulated at all, or to a lesser extent, at 30 °C. These include genes important for carbon and energy metabolism, as well as genes important for translation and metabolic activity; and these differences were observed from 2 hpi all the way to the end of the developmental cycle (released EBs). The evolved symbionts have a much lower infectivity than the ancestral ones, whether the infectivity assay was carried out at 20 °C or 30 °C. In fact, the evolved symbionts displayed even lower expression levels with much fewer genes being differentially expressed, and not even the T3SS genes were upregulated as expected at 2 hpi. So

right from the very start of the experiment, the chlamydiae were not primed for infection. This attenuation in infectivity levels continued throughout the developmental cycle, and thus, given this RNA-Seq data, the very low percent of infected amoebae at 96 hpi is not unexpected.

Conclusions

We initially set out to answer two broad questions. The first one asked whether simply carrying out the RNA-Seq experiment at an elevated temperature would affect gene expression dynamics. We answered this by comparing the previously obtained data on ancestral symbionts at 20 °C (König et al. 2017) to the data we obtained using ancestral symbionts at 30 °C. Although we did identify similarities in certain pathways, particularly in the upregulation of T3SS genes at 2 hpi, it does appear that there are different transcriptional dynamics. This is especially true later on in the developmental cycle, when the symbionts would have experienced the elevated temperature for a prolonged period of time. The elevated incubation temperature caused an attenuation in transcription levels, resulting in shallower than usual expression dynamics.

The second question sought to understand whether evolution at 30 °C brought about adaptation on a transcriptional level. To answer this we compared our two data sets obtained from the RNA-Seq experiments performed at 30 °C, on the ancestral and the evolved symbionts. We identified altered expression dynamics again, with the symbionts originating from the evolution experiment at 30 °C displaying even shallower expression dynamics. This was already evident 2 hpi and continued throughout the duration of the RNA-Seq experiment. Survival at this higher temperature may have required the symbionts to lower their expression levels, which in turn may have resulted in the reduction in infectivity as observed in our infectivity assays.

It would be interesting to assess whether the symbionts would regain their original expression levels if they were placed at the lower temperature once more and, if so, how long this would take to happen. In nature, one would expect periodic oscillations in temperature of even greater extremities. Winters can reach sub-zero temperatures, whereas summers can reach 40 °C. However, these symbionts may also survive as cysts to withstand such extreme conditions.

Materials and Methods

Evolution experiment

The evolution experiment was founded when ancestral *A. castellanii* Neff was freshly infected with *P. amoebophila* UWE25 (Collingro, Toenshoff, et al. 2005). After a continuous culture was established, the experiment was set up in 24-well plates (Thermo Fisher Scientific Inc., Waltham, MA, USA) under two temperatures (20 °C and 30 °C) and two bottleneck (BN) conditions (10% and 0.1%). Each treatment had 12 replicates with 1 ml TSY (30 g/l trypticase soy broth, Oxoid, 10 g/l yeast extract; pH 7.3) in each well. Populations were propagated by weekly 10-fold dilutions in 1 ml TSY in the 10% BN condition (allowing for ~3.35 generations per week), whereas monthly 1000-fold dilutions in 1 ml TSY were carried out in the 0.1% BN condition (allowing for ~11 generations per month). Whole-population samples were taken every 2-3 months and stored with DMSO as a cryoprotectant in liquid nitrogen, where they were available for any genome and phenotype analyses later on.

Infectivity assays

Protochlamydia EBs were freshly purified from *Acanthamoeba* cultures grown in 175 cm^2 culture flasks, in which EBs had been allowed to accumulate in TSY for 1 week. Purification of EBs was conducted as described previously. Symbiont-free amoebae were harvested right before infection and an aliquot was counted in a Neubauer chamber. The amoeba cell suspension in TSY was seeded in the required amount of wells (1 ml/well) in a 24-well plate at a concentration of 10^5 amoebae/ml. The plate was then incubated at 20 °C for 1 hour to allow the amoebae to attach prior to infection. EBs for the samples to be tested were thawed at 37 °C in a water bath and added in triplicate to wells in the 24-well plate at a multiplicity of infection (MOI) of 30, as this was predetermined to be the best MOI to observe differences in infectivity between samples. The well-plates were then incubated for 15 min at room temperature, followed by centrifugation (1,000 rpm, 15 min, 20 °C) so as to synchronize the infection. The medium was then changed so as to remove any residual EBs. This was done by removing 1 ml TSY per well by pipetting, and then re-adding fresh 1 ml TSY. Wells were then sampled at 6 hours post infection (hpi) and 96 hpi, marking the initial uptake of EBs and the completion of one developmental cycle, respectively. One of the wells with amoebae was treated the

same way as the rest but left uninfected, and this served as a negative control to ensure that there was no infection prior to incubation.

FISH using a Cy3-labeled probe (E25-454, sequence GGA TGT TAG CCA GCT CAT; Thermo Fisher Scientific, Waltham, MA, USA) was then performed as described elsewhere (Poppert et al. 2002; Wagner et al. 2003; Sixt et al. 2013) on the sampled time points, and cells were also stained with DAPI for 3 min. Using an epifluorescence microscope, the number of infectious particles in 50 amoebae from each sample were counted and divided into 3 categories: uninfected, slightly infected (1-30 particles) and highly infected (>30 particles), and this allowed for the infectivity of that particular sample from the assay to be established.

DNA extraction and sequencing

Frozen amoeba-chlamydia cultures were thawed in a water bath at 35 °C for 2 min and then pipetted into 25 cm^2 culture flasks (Nalge Nunc International, Rochester, NY, USA) with 10 ml TSY to allow sufficient growth. After one week, these cultures were then upscaled into 175 cm^2 flasks (Biologix Group Ltd, Shandong, China) for another week so as to increase the quantity of EBs that could be isolated. EB purification from the supernatant was done as described previously. Isolation of genomic DNA was then carried out on these isolated EBs using a cetyl trimethylammonium bromide-based extraction method as in (Schmitz-Esser et al. 2010). Clones were not sequenced per replicate but instead pool sequencing of symbiont populations was performed. Library preparation and pool sequencing were performed at the Vienna BioCenter Core Facilities (VBCF) Next-Generation Sequencing (NGS) Unit (https://www.viennabiocenter.org/facilities/). Illumina Genome Analyser and HiSeq2000 instruments were used to generate paired-end reads of ~125 bases according to standard procedures, achieving the desirable coverage of 300- to 600-fold.

Sequence read processing and mapping

The quality filtering and trimming of the Illumina reads was done by Prinseq-lite (v0.19.5) (Schmieder and Edwards 2011) and Trimmomatic (v0.32) (Bolger et al. 2014) and briefly applied as follows. First, a sliding window with size 10 removed any bases with lower quality than 20 starting from the 3’ side by removing the read part containing the low-quality bases. Any reads shorter than 40 nucleotides were also

removed. We discarded any low quality read that had an average Phred score below 30. Only read pairs were kept. These reads were mapped against the reference genome sequence of *P. amoebophila* UWE25 (Horn et al. 2004) (NCBI RefSeq accession number NC_005861.2) using the Burrows-Wheeler Aligner (BWA) (v0.7.5a) (Li and Durbin 2009) with standard settings and stored as bam files. For conversions from sam to bam files and from bam to fastq files (as Cortex_var input), we used SAMtools (v0.1.18) (Li et al. 2009) and Picard Tools (v1.92, http://broadinstitute.github.io/picard/).

Variant prediction

We used the VarCap pipeline (https://github.com/ma2o/VarCap) (Zojer et al. 2017) to detect single nucleotide and structural differences at high resolution and accuracy based on Illumina reads mapped to the reference *P. amoebophila* genome sequence. VarCap automatically performs quality filtering, mapping, variant calling and post-filtering of the predicted variants. The output is a Variant Call Format (VCF) file with a detailed description of the variants and their relative frequencies as well as two PDF files, which give a graphical overview of variant coverage and their frequency distribution.

Zojer and colleagues separately evaluated numerous variant detection tools for SNPs, insertion and deletions, as well as structural variants. Owing to the different variant calling abilities of the separate tools at low frequencies, several tools were combined so as to increase the sensitivity of the VarCap pipeline. Furthermore, a variant call had to be supported by at least two different tools so as to gain precision and robustness. Combining all these selected software tools, all variants except inversions could be detected at a minimum read abundance of 2% at 400x total coverage, with a minimum of 8 reads per variant, with high sensitivity (Zojer et al. 2017).

RNA-Seq infection experiments

EBs were freshly purified from amoeba cultures grown in 500 cm^2 culture flasks (Nalge Nunc International, Rochester, NY, USA), in which EBs had been allowed to accumulate in the medium for 1 week. These EBs were isolated from the ancestor

and from symbionts that had evolved for 510 generations at 30 °C . Purification of EBs and subsequent quantification of purified EBs was carried out as described previously. Symbiont-free amoebae were harvested 3 days before infection and seeded at low cell density, followed by incubation at 30 °C until infection. To optimize infection efficiency, particularly at early time points, we used an MOI of 145 for 2 hpi, an MOI of 80 for 48 hpi and an MOI of 30 for released EBs accumulated at 7 days post infection. Precultivated amoebae were harvested, transferred to 50 ml Greiner tubes (Greiner Bio-One GmbH, Frickenhausen, Germany), and purified *Protochlamydia* EBs were added, followed by repeated centrifugation (centrifuged at 130 x g twice for 5 min each time and then once for 10 min at 30 °C) with vortexing between the centrifugation steps. Infected amoebae were then transferred back to the culture flasks and incubated in TSY medium at 30 °C for 2 hr before the infection was synchronised by gently washing the attached amoebae three times with PAS. TSY medium was added to the cultures; some culture flasks were sampled at the 2 hpi time point, whereas the remaining culture flasks were incubated at 30 °C for 48 hr. Released *Protochlamydia* EBs were harvested from the medium supernatant after one week of incubation from amoeba-endosymbiont cultures that were grown at the same conditions, and were purified as described previously. All infection experiments were performed in biological triplicates.

RNA extraction and sequencing

In order to increase the coverage of the *Protochlamydia* transcriptome, a protocol for enrichment of bacteria prior to RNA extraction was employed (König et al. 2017). Each sample was processed in less than 7 min so as to minimize possible changes of the transcriptomes during enrichment, since the half-life of total mRNA from *E. coli* was demonstrated to be in this range (Selinger et al. 2003). Infected amoebae were harvested and collected (7,600 X g, 2 min, 30 °C), after which the pellets were resuspended in a sucrose buffer (35 mM Tris-HCl, 250 mM sucrose, 25 mM KCl, 10 mM $MgCl_2$) supplemented with 50 *µ*g/ml rifampicin (Sigma-Aldrich Handels GmbH, Vienna, Austria) in order to inhibit active transcription during the enrichment procedure (Sarov and Becker 1971; Dreses-Werringloer et al. 2001). Amoebae were then disrupted by vortexing in the presence of glass beads (diameter of 0.75 to 1 mm; Carl Roth) for 1 min. The suspensions were subsequently filtered through a 5

*μ*m filter, the flowthrough fractions containing the bacteria were collected by centrifugation (10,600 X g, 2 min, room temperature), and the pellets were immediately resuspended in TRIzol reagent (Thermo Fisher Scientific). Extracellular *Protochlamydia* EBs were pelleted (20,800 X g, 2 min, room temperature); these pellets were then resuspended in sucrose buffer and treated like the enriched bacteria.

Cells were mechanically disrupted by bead beating for 30 s at 4.5 m/s using lysing matrix A tubes and a FastPrep-24 instrument (MP Biomedicals, Santa Ana, CA, USA). Subsequent RNA extraction was performed according to the TRIzol guidelines. Residual DNA was then digested using the Turbo DNA-free kit (Thermo Fisher Scientific) as recommended by the manufacturer. DNase-treated RNA was precipitated with sodium acetate and ethanol and dissolved in nuclease-free water (Thermo Fisher Scientific), whereas DNA contamination was controlled for via PCR targeting a short region of the bacterial 16S rRNA gene (SigF2/R2 primers; Haider et al. 2008) using 35 PCR cycles. rRNA was subsequently removed using the Ribo-Zero magnetic kit for Gram-positive bacteria according to the manufacturer's instructions (Illumina, San Diego, CA, USA). For strand-specific cDNA library preparation, the NEBNext Ultra directional RNA library prep kit for Illumina, in combination with the NEBNext multiplex oligonucleotides (New England Biolabs, Ipswich, MA, USA), was used starting at first-strand cDNA synthesis. All libraries were sequenced using an Illumina HiSeq2500 system at the VBCF NGS Unit with 50 bp read length.

RNA-Seq read processing

Sequencing reads were trimmed and cleaned before mapping, based on sequencing read statistics obtained using Trimmomatic (v0.32) (Bolger et al. 2014) and FastQC (v0.10.0) (Andrews 2010). Briefly, (i) the first 11 bases were removed, (ii) any remaining adapter sequences were clipped, and (iii) sequence reads shorter than 25 bases were removed, all done using Trimmomatic. To map bacterial reads to the *Protochlamydia amoebophila* reference genome (Horn et al. 2004) BWA (v0.7.5a) (Li and Durbin 2009) was used. Only unambiguously mapped reads were kept using SAMtools (v0.1.18) (Li et al. 2009). Strand-specific reads per predicted gene were counted via HTSeq (v0.11.1) (Anders et al. 2015).

Gene expression analyses

Differentially expressed genes were determined between two consecutive time points (2 hpi to 48 hpi, 48 hpi to EBs, EBs to 2 hpi) using the R software environment (v3.6.1) and the Bioconductor package DESeq2 (v1.22.1) (Gentleman et al. 2004; R Development Core Team. 2011; Love et al. 2014). Genes were considered differentially expressed if their expression changed twofold with a false-discovery rate (FDR) smaller or equal 0.05. Log_2 fold changes (logFC) of differentially expressed genes and mean centered gene expression values (log_2 transcripts per kilobase million [TPM]) were used for visualization as heat maps using the Rpackage gplots (v3.0.1.1) (Warnes et al. 2005).

To extend and improve the available *Protochlamydia* genome annotation by Horn et al. 2004 for each gene, we collected Pfam domains (Bateman et al. 2002), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and gene ontology (GO) term level 5 assignments using DAVID (Huang et al. 2009), COG (clusters of orthologous groups of proteins) cluster and class assignments using MaGe (Vallenet 2006), and type III secretion effector predictions using Effective (Jehl et al. 2011). The Bioconductor software package GOseq (Young et al. 2010) was used to test for statistical enrichment of functional categories (FDR $\leq$ 0.05) among different gene sets. To test whether predicted type III secreted proteins were significantly enriched in any given gene set, two-tailed Fisher's exact tests were conducted, and values below 0.05 were considered statistically significant.

Statistics

Statistical analysis was performed using R (v3.6.1) (R Development Core Team. 2011). Details on sample sizes, in addition to the statistical tests conducted, are shown in the corresponding figure legends; the statistical parameters and significance are reported in the figure legends. Data were considered to be statistically significant when $P \leq 0.05$. Comparisons for two groups were calculated by a two-sided Wilcoxon–Mann–Whitney test. Comparisons for paired samples were calculated by a Wilcoxon signed-rank test.

References

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. *FastQC* [Internet]. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL. 2002. The Pfam protein families database. *Nucleic Acids Res.* [Internet] 30:276 –280. Available from: https://doi.org/10.1093/nar/30.1.276

Belland RJ, Zhong G, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL, Caldwell HD. 2003. Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc. Natl. Acad. Sci. U. S. A*. 100:8478–8483.

Betts-Hampikian HJ, Fields KA. 2010. The Chlamydial Type III Secretion Mechanism: Revealing Cracks in a Tough Nut. *Front. Microbiol.* [Internet] 1. Available from: http://dx.doi.org/10.3389/fmicb.2010.00114

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.

Chiner-Oms Á, Sánchez-Busó L, Corander J, Gagneux S, Harris SR, Young D, González-Candelas F, Comas I. 2019. Genomic determinants of speciation and spread of the *Mycobacterium tuberculosis* complex. *Sci. Adv.* 5:eaaw3307.

Collingro A, Poppert S, Heinz E, Schmitz-Esser S, Essig A, Schweikert M, Wagner M, Horn M. 2005. Recovery of an environmental *Chlamydia* strain from activated sludge by co-cultivation with *Acanthamoeba* sp. *Microbiology* 151:301–309.

Collingro A, Toenshoff ER, Taylor MW, Fritsche TR, Wagner M, Horn M. 2005. "*Candidatus* Protochlamydia amoebophila", an endosymbiont of *Acanthamoeba* spp. *Int. J. Syst. Evol. Microbiol.* 55:1863–1866.

Crépin S, Chekabab S-M, Le Bihan G, Bertrand N, Dozois CM, Harel J. 2011. The Pho regulon and the pathogenesis of *Escherichia coli*. *Vet. Microbiol.* [Internet] 153:82–88. Available from: http://dx.doi.org/10.1016/j.vetmic.2011.05.043

Dreses-Werringloer U, Padubrin I, Zeidler H, Köhler L. 2001. Effects of azithromycin and rifampin on *Chlamydia trachomatis* infection in vitro. *Antimicrob. Agents Chemother.* 45:3001–3008.

Fawcett P, Eichenberger P, Losick R, Youngman P. 2000. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* 97:8063–8068.

Haferkamp I, Schmitz-Esser S, Wagner M, Neigel N, Horn M, Neuhaus HE. 2006. Tapping the nucleotide pool of the host: novel nucleotide carrier proteins of *Protochlamydia amoebophila*. *Mol. Microbiol.* 60:1534–1545.

Haider S, Collingro A, Walochnik J, Wagner M, Horn M. 2008. *Chlamydia*-like bacteria in respiratory samples of community-acquired pneumonia patients. *FEMS Microbiol. Lett.* 281:198–202.

Herrera P, Schuster S, Wentrup C, König L, Kempinger T, Na H, Schwarz J, Köstlbacher S, Wascher F, Zojer M, Rattei T, Horn M. 2020. Molecular causes of an evolutionary shift along the parasitism-mutualism continuum in a bacterial symbiont. *Proc. Natl. Acad. Sci. U. S. A.* doi:10.1073/pnas.2005536117.

Herrera P, Schuster S, Zojer M, Na H, Schwarz J, Wascher F, Kempinger T, Regner A, Rattei T, Horn M. 2020. Genome dynamics and temperature adaptation during experimental evolution of obligate intracellular bacteria. Manuscript in preparation.

Horn M. 2008. *Chlamydiae* as symbionts in eukaryotes. *Annu. Rev. Microbiol.* 62:113–131.

Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B, Brandt P, Nyakatura GJ, Droege M, Frishman D, et al. 2004. Illuminating the evolutionary history of Chlamydiae. *Science* 304:728–730.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4:44–57.

Jehl M-A, Arnold R, Rattei T. 2011. Effective—a database of predicted secreted bacterial proteins. *Nucleic Acids Res.* 39:D591–D595.

Johnson DL, Stone CB, Bulir DC, Coombes BK, Mahony JB. 2009. A novel inhibitor of *Chlamydophila pneumoniae* protein kinase D (PknD) inhibits phosphorylation of CdsD and suppresses bacterial replication. *BMC Microbiol.* 9:218.

König L, Siegl A, Penz T, Haider S, Wentrup C, Polzin J, Mann E, Schmitz-Esser S, Domman D, Horn M. 2017. Biphasic Metabolism and Host Interaction of a Chlamydial Symbiont. *mSystems* [Internet] 2. Available from: http://dx.doi.org/10.1128/mSystems.00202-16

Konkel ME, Tilly K. 2000. Temperature-regulated expression of bacterial virulence genes. *Microbes Infect.* 2:157–166.

Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L. 2000. Global analysis of the genetic network controlling a bacterial cell cycle. *Science* 290:2144–2148.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.

Mäurer AP, Mehlitz A, Mollenkopf HJ, Meyer TF. 2007. Gene expression profiles of *Chlamydophila pneumoniae* during the developmental cycle and iron

depletion-mediated persistence. *PLoS Pathog.* 3:e83.

Maurin M, Bryskier A, Raoult D. 2002. Antibiotic susceptibilities of *Parachlamydia acanthamoeba* in amoebae. *Antimicrob. Agents Chemother.* 46:3065–3067.

Poppert S, Essig A, Marre R, Wagner M, Horn M. 2002. Detection and differentiation of chlamydiae by fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 68:4081–4089.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robert C Gentleman, Vincent J Carey, Douglas M Bates, Ben Bolstad, Marcel Dettling, Sandrine Dudoit, Byron Ellis, Laurent Gautier, Yongchao Ge, Jeff Gentry, Kurt Hornik, Torsten Hothorn, Wolfgang Huber, Stefano Iacus, Rafael Irizarry, Friedrich Leisch, Cheng Li, Martin Maechler, Anthony J Rossini, Gunther Sawitzki, Colin Smith, Gordon Smyth, Luke Tierney, Jean YH Yang & Jianhua Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* [Internet] 5. Available from: https://genomebiology.biomedcentral.com/articles/10.1186/gb-2004-5-10-r80

Sarov I, Becker Y. 1971. Deoxyribonucleic acid-dependent ribonucleic acid polymerase activity in purified trachoma elementary bodies: effect of sodium chloride on ribonucleic acid transcription. *J. Bacteriol.* 107:593–598.

Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864.

Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wagner M, Rattei T, Horn M. 2010. The genome of the amoeba symbiont "*Candidatus* Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* 192:1045–1057.

Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C. 2003. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* 13:216–223.

Sixt BS, Hiess B, König L, Horn M. 2012. Lack of effective anti-apoptotic activities restricts growth of *Parachlamydiaceae* in insect cells. *PLoS One* 7:e29565.

Sixt BS, Siegl A, Müller C, Watzka M, Wultsch A, Tziotis D, Montanaro J, Richter A, Schmitt-Kopplin P, Horn M. 2013. Metabolic features of *Protochlamydia amoebophila* elementary bodies--a link between activity and infectivity in *Chlamydiae*. *PLoS Pathog.* 9:e1003553.

Subtil A, Collingro A, Horn M. 2014. Tracing the primordial *Chlamydiae*: extinct parasites of plants? *Trends Plant Sci.* 19:36–43.

Tjaden J, Winkler HH, Schwöppe C, Van Der Laan M, Möhlmann T, Neuhaus HE. 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.

Vallenet D. 2006. MaGe: a microbial genome annotation system supported by

synteny results. *Nucleic Acids Res.* [Internet] 34:53–65. Available from: http://dx.doi.org/10.1093/nar/gkj406

Wagner M, Horn M, Daims H. 2003. Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Curr. Opin. Microbiol.* 6:302–309.

Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S. 2005. gplots: various R programming tools for plotting data. R package version 2. R Foundation for Statistical Computing, Vienna, Austria.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14.

Zojer M, Schuster LN, Schulz F, Pfundner A, Horn M, Rattei T. 2017. Variant profiling of evolving prokaryotic populations. *PeerJ* 5:e2997.

Survival at an elevated temperature strongly reduces gene expression dynamics in a chlamydial symbiont

Paul Herrera, Cecilia Wentrup, Lena König, Stephan Köstlbacher, Thomas Rattei, Matthias Horn

Supplementary information

Supplementary Table 3: Summary of all novel variants identified in replicate 5 at 30 ^{o}C in the evolution experiment.

Structural variant type	Variant effect	NCBI locus tag	Gene name	Freq (%)	RefSeq function
SNP	nonsynonymous	PC_RS02320	ptsI	22.3	phosphoenolpyruvate-protein phosphotransferase (PTS enzyme I)
SNP	nonsynonymous	PC_RS03870		35.8	hypothetical protein
SNP	nonsynonymous	PC_RS07875	pknD	55.4	hypothetical protein
SNP	nonsynonymous	PC_RS01680		15.6	hypothetical protein
SNP	synonymous	PC_RS04675		3.0	hypothetical protein
SNP	nonsynonymous	PC_RS04675		2.8	hypothetical protein
SNP	nonsynonymous	PC_RS08925		20.6	MFS transporter
SNP	nonsynonymous	PC_RS09030	phoR	23.7	two-component sensor histidine kinase phoR
INS	frameshift	PC_RS05410		91.2	HlyC/CorC family transporter
DEL	frameshift	PC_RS00480		22.3	inorganic phosphate transporter
DEL	frameshift	PC_RS06415		27.1	ABC transporter ATP-binding protein
INS	frameshift	PC_RS06415		22.3	ABC transporter ATP-binding protein
INS	frameshift	PC_RS04995		93.7	23S ribosomal RNA

Chapter VI

Concluding discussion

Conclusion

Although *Chlamydiae* have long been considered to be highly successful bacterial pathogens of humans and animals, the relatively recent discovery of the environmental chlamydiae in free-living amoebae (Amann et al. 1997; Fritsche et al. 2000; Horn et al. 2000) found ubiquitously in nature has radically changed our perception of this remarkable group of bacteria. This subsequently led to numerous open questions relating to the evolution of these bacteria, their intracellular lifestyle and the different ways they can adapt to their eukaryotic hosts. The principal goal of my work here was to provide novel insights into the evolutionary dynamics of these symbionts in their amoeba hosts, making use of long-term evolution experiments that are generally limited to free-living microbes. I combined this approach with infection assays, genome re-sequencing, and global gene expression analysis to explore the different routes that adaptation can take. I sought to understand how different members of the environmental chlamydiae can adapt and overcome challenges presented to them in their environment, namely imposed transmission mode and elevated temperature.

The study presented in the third chapter of this thesis explores the fact that *Chlamydiae* can be vertically transmitted from parent to daughter cells during host cell division, but can also readily infect naive hosts. We made use of this mixed transmission mode (Ebert 2013) that is inherent to our bacteria to experimentally manipulate symbiont transmission among host generations. Similar to previous experimental studies (Bull et al. 1991; Stewart et al. 2005; Sachs and Wilcox 2006), we also demonstrated that horizontally transmitted symbionts are more parasitic whereas vertically transmitted symbionts are more mutualistic. Pool sequencing of bacteria then allowed us to assess any genomic changes that had occurred. We discovered that mutations arising during the experiment were relatively few and most remained at low frequencies. On the other hand, we observed a strong shift in ancestral variant frequencies under conditions of horizontal transmission. This resulted in a major change in the symbiont population concomitant with the increase in parasitism during the experiment. We also performed RNA sequencing analysis on symbionts from the two different transmission mode regimes, and we demonstrated that the gene expression profiles differed at all time points analyzed

between the two regimes, indicating marked changes in temporal gene regulation. An exciting revelation was that the increased parasitism was a result of processes occurring at the infectious stage of the symbiont's developmental cycle. Our multilevel analysis provides the first comprehensive picture of how a bacterial symbiont can shift its relationship with its host between parasitic and mutualistic.

The molecular data presented in this study is hopefully only the beginning of a new wave of studies on this topic. It would be intriguing to understand how other symbioses adapt on the molecular level when their mode of transmission is altered. Would adaptations be more similar among different bacteria but vary when compared to viruses? Or would the type of host be a more important contributing factor? Although we showed that a shift in ancestral variants was linked to the increased parasitism, other systems may be more dependent on beneficial mutations that arise during the experiment. We revealed that genes involved in energy production, required for extracellular survival, and the symbiont's primary virulence mechanism were significantly upregulated for horizontally transmitted bacteria. This is directly linked to the unique developmental cycle of the *Chlamydiae*. Thus, exploring the gene expression dynamics of other symbionts may reveal whether there are common features among different systems, or whether each system has a unique way of modifying their gene expression profiles. Ultimately, these studies together will paint a clearer picture of how parasites can emerge from the plethora of known symbiotic interactions when their transmission mode is altered.

Our evolution experiment spanning more than three years sought to take a closer look at genome evolution of these intracellular bacteria in a way that is normally restricted to free-living microbes. I go into great depth about this study in the fourth chapter of the thesis, where we analyzed the evolved phenotypes and genome dynamics over the course of the experiment. By making use of different dilution and temperature conditions, we could elucidate that adaptation varied between treatments, in particular when comparing ambient to elevated temperature. Although replicates from both temperatures adapted to their growth environment, there was stronger evidence for positive selection at 30 °C, which had been expected to be a more challenging temperature for both host and symbiont. We observed a significant increase in the number of nonsynonymous substitutions relative to synonymous

ones, as well as an overrepresentation of small indels in the symbionts isolated from the 10% dilution at 30 °C treatment. Apart from this, mutations in this treatment were more likely to persist over multiple time points and increase to high frequencies. Taken together, these results suggest that numerous mutations at the elevated temperature were beneficial and allowed the symbiont to adapt to this stressful environment. This genomic adaptation was coupled with attenuation of symbiont infectivity potentially as a trade-off to minimise host cell burden.

Although the findings from our study provide initial insights into how these unique bacteria can adapt to their surrounding environment, so much more can still be explored. By sequencing more replicates and extending the timeline of this evolution experiment we would be able to assess how genetic parallelism changes over time and whether novel phenotypes occur. The experiment could also be gradually shifted to even higher temperatures and this could provide more answers relating to how these host-dependent symbionts made the leap from protists to become major pathogens of humans. With techniques such as fluorescence-activated cell sorting we could even be able to analyze the symbiont population in one single amoeba cell and therefore understand the dynamics of adaptation in this structured environment.

The study here was focused on only one chlamydial species and so this work opens the door to various follow-up questions on other members of the environmental chlamydiae. Do they share common adaptive routes? And if they differ, would this depend more strongly on the host they infect or perhaps on the environment they are found in? The virulence of the species could also play a crucial role in evolutionary dynamics since some environmental chlamydiae are not so benign and can even cause disease in humans (Friedman et al. 2003; Collingro et al. 2005; Baud et al. 2008). Apart from related environmental chlamydiae, such long-term experiments should definitely be extended to other host-dependent systems where similar questions can be asked about how such microorganisms adapt on different levels. Only by increasing the amount of research carried out on diverse species can we truly make broad statements about how different forms of life adapt and overcome challenges in their surrounding environment.

The fifth chapter of this thesis is an extension of the evolution experiment by adding RNA sequencing data to the already acquired phenotypic and genomic data. The aim of this study was to understand how transcriptional dynamics were affected by prolonged growth at 30 °C. By comparing the data we obtained at 30 °C to previous work carried out in our lab at 20 °C (König et al. 2017), it was evident that the elevated temperature resulted in altered expression dynamics. More specifically, gene expression dynamics were strongly reduced both due to elevated incubation temperature and especially as a result of long-term cultivation at an elevated temperature. Far fewer genes were differentially expressed in the 30 °C evolved symbionts throughout the experiment, and this was observed already at the earliest time point. Pathways that are essential for proper infection and development were not upregulated, most importantly those involved in virulence and carbon and energy metabolism. These transcriptional dynamics suggest that the symbionts maintained at 30 °C were not primed for infection, and could have resulted in the observed attenuation in infectivity. The work presented in this chapter thus adds an extra level to our understanding of the way these bacteria can adapt to their environment.

In summary, this thesis has provided an initial glimpse at some of the adaptive processes occurring in a host-dependent symbiont. Such symbiotic relationships with microbes are ubiquitous in nature and present in practically all eukaryotes, shaping the world around us. Understanding how these intracellular bacteria overcome novel challenging conditions is important in today's world, both due to current environmental issues such as climate change, but also due to the neverending arms race between our immune system and pathogens. By studying a diversity of these systems we may elucidate common patterns of how symbionts are able to transition from benign mutualists to lethal pathogens. This is also true for the environmental chlamydiae due to the ability of some members to cause disease in animals and humans. Another aspect that tends to be overlooked is the importance of protists in soil fertility and nutrient recycling. These protists also feed on soil bacteria and release nutrients in the process. However, environmental chlamydiae are able to evade digestion and co-exist with their protist hosts, even providing protection from pathogens that can infect humans (König et al. 2019). More research invested into how these bacteria survive and adapt to their environment may therefore not only be of basic research interest but also of great significance to public health.

References

Amann R, Springer N, Schönhuber W, Ludwig W, Schmid EN, Müller KD, Michel R. 1997. Obligate intracellular bacterial parasites of acanthamoebae related to *Chlamydia* spp. *Appl. Environ. Microbiol.* 63:115–121.

Baud D, Regan L, Greub G. 2008. Emerging role of *Chlamydia* and *Chlamydia*-like organisms in adverse pregnancy outcomes. *Curr. Opin. Infect. Dis*. 21:70–76.

Bull JJ, Molineux IJ, Rice WR. 1991. Selection of Benevolence in a Host-Parasite System. *Evolution* [Internet] 45:875. Available from: http://dx.doi.org/10.2307/2409695

Collingro A, Poppert S, Heinz E, Schmitz-Esser S, Essig A, Schweikert M, Wagner M, Horn M. 2005. Recovery of an environmental *Chlamydia* strain from activated sludge by co-cultivation with *Acanthamoeba* sp. *Microbiology* 151:301–309.

Ebert D. 2013. The Epidemiology and Evolution of Symbionts with Mixed-Mode Transmission. *Annu. Rev. Ecol. Evol. S.* [Internet] 44:623–643. Available from: http://dx.doi.org/10.1146/annurev-ecolsys-032513-100555

Friedman MG, Dvoskin B, Kahane S. 2003. Infections with the chlamydia-like microorganism *Simkania negevensis*, a possible emerging pathogen. *Microbes Infect.* 5:1013–1021.

Fritsche TR, Horn M, Wagner M, Herwig RP, Schleifer KH, Gautom RK. 2000. Phylogenetic diversity among geographically dispersed *Chlamydiales* endosymbionts recovered from clinical and environmental isolates of *Acanthamoeba* spp. *Appl. Environ. Microbiol.* 66:2613–2619.

Horn M, Wagner M, Müller K-D, Schmid EN, Fritsche TR, Schleifer K-H, Michel R. 2000. *Neochlamydia hartmannellae* gen. nov., sp. nov. (*Parachlamydiaceae*), an endoparasite of the amoeba *Hartmannella vermiformis*. *Microbiology* 146 (Pt 5):1231–1239.

König L, Siegl A, Penz T, Haider S, Wentrup C, Polzin J, Mann E, Schmitz-Esser S, Domman D, Horn M. 2017. Biphasic Metabolism and Host Interaction of a Chlamydial Symbiont. *mSystems* [Internet] 2. Available from: http://dx.doi.org/10.1128/mSystems.00202-16

König L, Wentrup C, Schulz F, Wascher F, Escola S, Swanson MS, Buchrieser C, Horn M. 2019. Symbiont-Mediated Defense against *Legionella pneumophila* in Amoebae. *MBio* [Internet] 10. Available from: http://dx.doi.org/10.1128/mBio.00333-19

Sachs JL, Wilcox TP. 2006. A shift to parasitism in the jellyfish symbiont *Symbiodinium microadriaticum*. *Proc. Biol. Sci.* 273:425–429.

Stewart AD, Logsdon JM Jr, Kelley SE. 2005. An empirical study of the evolution of virulence under both horizontal and vertical transmission. *Evolution* 59:730–739.

Appendix

Abstract / Zusammenfassung

Abstract

The *Chlamydiae* are an ancient group of obligate intracellular bacteria that reside within eukaryotic cells and display a unique biphasic developmental cycle. They are of global significance because of their success as major human pathogens as well as causing numerous diseases in a variety of animals. But over the past few decades we have discovered an enormous diversity of *Chlamydiae* that are associated with an impressive range of eukaryotic hosts. The majority of these so-called environmental chlamydiae are considered to be symbionts of unicellular eukaryotes—free-living amoebae or other protists—that are found ubiquitously in the environment. The main aim of my research was to improve our understanding of the evolution of these unique host-dependent bacteria living in amoebae. I combined experimental evolution approaches, infectivity assays, genome re-sequencing, and global gene expression analysis to achieve this goal. The first study involved the experimental manipulation of the transmission of symbionts among host generations to favour horizontal or vertical transmission using *Parachlamydia acanthamoebae*. We revealed that increased parasitism in the horizontally transmitted symbionts was driven by a dramatic shift in the frequency of standing genetic variants coupled with an extensive modification of gene expression. In particular, we observed an upregulation of genes affecting functions crucial at the infectious stage of the symbiont and required for extracellular survival. We then used *Protochlamydia amoebophila* for the rest of our work as a model to gain initial insight into the process of temperature adaptation in these symbionts, a crucial first step in the evolution of the major chlamydial pathogens. We showed evidence for positive selection at the elevated temperature and observed an attenuation in symbiont infectivity, very likely so as to reduce host cell burden. In addition, shallow expression dynamics throughout the infection cycle may have contributed to this reduced infectivity. Taken together, our work has provided some fascinating insights into the evolution and adaptation of these remarkable chlamydial symbionts in their protected intracellular niche.

Zusammenfassung

Chlamydien sind eine evolutionsgeschichtlich sehr ursprüngliche Gruppe obligat intrazellulärer Bakterien, die in eukaryotischen Zellen leben und einen einzigartigen biphasischen Entwicklungszyklus aufweisen. Weltweit sind sie bedeutende Krankheitserreger des Menschen und Auslöser zahlreicher Erkrankungen bei einer Vielzahl von Tieren. In den letzten Jahrzehnten wurden eine enorme Diversität an Chlamydien entdeckt, die mit einer beeindruckenden Vielfalt eukaryotischer Wirte assoziiert sind. Viele dieser so genannten Umweltchlamydien leben als Symbionten einzelliger Eukaryonten, wie Amöben oder andere Protisten, die in der Umwelt allgegenwärtig sind. Das Hauptziel meiner Forschung war es, die Evolution dieser einzigartigen wirtsabhängigen Bakterien besser zu verstehen. Hierzu habe ich verschiedene Evolutionsexperimente mit Infektivitätsassays, Genom-Resequenzierung und globaler Genexpressionsanalyse kombiniert. Im Rahmen der ersten Studie untersuchte ich den Einfluss der Übertragungswege der Symbionten zwischen den Wirtsgenerationen. Im Evolutionsexperiment wurden Bedingungen geschaffen, um die horizontale oder vertikale Übertragung für *Parachlamydia acanthamoebae* zu begünstigen. Wir konnten zeigen, dass erhöhter Parasitismus bei den horizontal übertragenen Symbionten durch eine dramatische Verschiebung der Populationszusammensetzung in Verbindung mit einer umfassenden Veränderung der Genexpression hervorgerufen wird. Insbesondere beobachteten wir eine Hochregulation von Genen, die Funktionen beeinflussen, die im infektiösen Stadium des Symbionten entscheidend und für das extrazelluläre Überleben notwendig sind. Für den Rest unserer Arbeit verwendeten wir *Protochlamydia amoebophila* als Modell, um einen ersten Einblick in den Prozess der Temperaturanpassung in diesen Symbionten zu gewinnen, ein entscheidender Schritt in der Evolution der wichtigsten Chlamydien, die Tiere infizieren. Wir konnten positive Selektion bei erhöhter Temperatur nachweisen und beobachteten eine Abschwächung der Infektiosität der Symbionten im Verlauf des Evolutionsexperiments, was sehr wahrscheinlich die Belastung der Wirtszellen verringert. Darüber hinaus könnte eine verminderte Expressionsdynamik während des gesamten Infektionszyklus zu dieser verringerten Infektiosität beigetragen haben. Zusammengenommen hat unsere Arbeit faszinierende Einblicke in die

Evolution und Anpassung dieser bemerkenswerten chlamydialen Symbionten in ihrer geschützten intrazellulären Nische geliefert.