BIOZENTRUM
Universität Basel
The Center for Molecular Life Sciences

PhD Thesis Benjamin Steeb

A quantitative analysis of *Salmonella* Typhimurium metabolism during infection

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Benjamin Steeb

aus Bad Kreuznach, Deutschland

Basel, 2012

creative commons

Namensnennung-Keine kommerzielle Nutzung-Keine Bearbeitung 2.5 Schweiz

---

Sie dürfen:

das Werk vervielfältigen, verbreiten und öffentlich zugänglich machen

Zu den folgenden Bedingungen:

Namensnennung. Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen (wodurch aber nicht der Eindruck entstehen darf, Sie oder die Nutzung des Werkes durch Sie würden entlohnt).

Keine kommerzielle Nutzung. Dieses Werk darf nicht für kommerzielle Zwecke verwendet werden.

Keine Bearbeitung. Dieses Werk darf nicht bearbeitet oder in anderer Weise verändert werden.

- Im Falle einer Verbreitung müssen Sie anderen die Lizenzbedingungen, unter welche dieses Werk fällt, mitteilen. Am Einfachsten ist es, einen Link auf diese Seite einzubinden.
- Jede der vorgenannten Bedingungen kann aufgehoben werden, sofern Sie die Einwilligung des Rechteinhabers dazu erhalten.
- Diese Lizenz lässt die Urheberpersönlichkeitsrechte unberührt.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/ Datum: 3.4.2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

- Prof. Dr. Dirk Bumann
- Prof. Dr. Christoph Dehio

Basel, den 21. Februar 2012

Prof. Dr. Martin Spiess
Dekan

Humboldtstr. 8
79576 Weil am Rhein
Deutschland

Table of contents

Abstract:

In this thesis, *Salmonella* metabolism during infection was investigated. The goal was to gain a quantitative and comprehensive understanding of *Salmonella in vivo* nutrient supply, utilization and growth.

To achieve this goal, we used a combined experimental / *in silico* approach. First, we generated a reconstruction of *Salmonella* metabolism ([1], see 2.1). This reconstruction was then combined with *in vivo* data from experimental mutant phenotypes to build a comprehensive quantitative *in vivo* model of *Salmonella* metabolism during infection (unpublished data, see 2.2). The data indicated that *Salmonella* resided in a quantitatively nutrient poor environment, which limited *Salmonella in vivo* growth. On the other hand, the *in vivo* niche of *Salmonella* was qualitatively rich with at least 45 different metabolites available to *Salmonella*. We then used the *in vivo* model of infection to analyze reasons for the preponderance of *Salmonella* genes with undetectable virulence phenotype (unpublished data, see 2.3). Our data indicated that host supply with diverse nutrients resulted in large-scale inactivity of numerous *Salmonella* metabolic pathways. This together with extensive metabolic redundancy was the main cause of the massive *Salmonella* gene dispensability during infection. To verify this hypothesis experimentally, an unbiased method for large scale mutagenesis was developed (unpublished data, see 2.4). Results from 20 *Salmonella* mutator lines indicate that *Salmonella* can tolerate at least some 2700 to 3900 mutations, emphasizing again that a vast majority of *Salmonella* genes are dispensable in a defined environment.

1. Introduction

1.1 *Salmonella* morphology and phylogeny

Salmonella is a Gram-negative, rod shaped gammaproteobacterium [2]. The genus *Salmonella* is divided into two species, *S. bongori* and *S*. e*nterica* [3, 4]. Whereas *S*. *bongori* is associated with cold-blooded animals [5], *S*. *enterica* is a pathogen for warm-blooded animals. *S. enterica* is divided into six subspecies. These subspecies contain over 2500 different *Salmonella* strains (serovars), which differ in their host specificity and can be distinguished by distinct antibody responses against their outer membrane and their flagella [6, 7]. *Salmonella enterica* subspecies *enterica* (subspecies I) forms with 1531 serovars the biggest group and contains relevant human pathogens like the serovars Typhi, Enteritis and Typhimurium [7-9].

1.2 *Salmonella* epidemiology and pathology

Salmonella is best known for two diseases it causes in humans - gastroenteritis and typhoid fever, which are caused by different *Salmonella* serovars. Over 90 million persons are infected with gastroenteritis causing *Salmonella* per year (average mortality rate 0.17%) [10]. Compared to this, the annual infection rate with typhoid fever causing *Salmonella* is lower (21.6 million persons per year), but typhoid fever cases have a higher mortality rate (1%) [11].

In both diseases, infection with *Salmonella* starts with ingestion of contaminated food or water. *Salmonella* passes through the stomach into the gut. Non-typhoidal *Salmonella* serovars like *S*. Enteritis and *S*. Typhimurium replicate in healthy persons only in the intestine, which causes gastroenteritis. Disease symptoms include diarrhea, fever, vomiting and abdominal pain. The disease is normally self-limiting with symptoms resolving after five to seven days (reviewed in [12, 13]). Typhoid fever on the other hand is caused by the human specific pathogen *S.* Typhi. The pathogen can cross the intestinal barrier into the bloodstream, and disseminates into the whole organism (systemic disease). During the course of the disease, bacteria can be found in spleen, liver, bone marrow and the gall bladder. Hallmarks of this disease are prolonged and

progressively rising fever, nausea, muscle pain and bronchitic cough, which can take 3 to 4 weeks to resolve [14, 15]. In 3 – 5% of the cases, typhoid fever develops into an asymptomatic chronic infection. There, *S.* Typhi colonizes the gall bladder and is excreted via the feces for weeks up to decades. Chronically infected individuals are healthy carriers that form reservoirs for the pathogen. The most famous example of a chronic carrier was the cook Mary Mallone ("Typhoid Mary"). Before she was imprisoned in 1907, she infected within seven years around 57 persons with *S.* Typhi ([16, 17], reviewed in [18]). In the elderly, children or persons with immune system deficiencies, non-typhoidal *Salmonella* (NTS) can also cause a systemic infection. NTS cases show symptoms similar to typhoid fever and have a high mortality rate (20%) ([19], reviewed in [20]).

1.3 Model systems to analyze *Salmonella* infections

S. Typhi causes severe infections in humans, but a practical model to analyze this infection in small animals is still lacking. *S.* Typhi is adapted to humans and does not cause infections in normal mice. On the other hand, mice with a humanized immune system can be successfully infected with *S.* Typhi. However, these experiments are labour intensive and there are big variations in the course of infection and *Salmonella* load between individual animals [21, 22].

Current research is therefore focused on models with non-typhoidal strains to analyze *Salmonella* virulence traits and processes. The model used most frequently is infection of genetically susceptible Balb/C mice with *S.* Typhimurium. Balb/C mice lack the divalent cation transporter Slc11a1, which is essential for control of *Salmonella* infection [23-25]. Infection of Balb/C mice with *S.* Typhimurium leads to a systemic disease similar to typhoid fever in humans. This includes *Salmonella* systemic spread to bone marrow, spleen and liver and replication in macrophages. *S.* Typhimurium infection in susceptible mice also leads to an enlargement of spleen and liver (hepatosplenomegaly) and the encapsulation of the pathogens by clusters of phagocytic cells (granulomas) in these tissues ([26-28], reviewed in [13]).

1.4 Pathogenesis of *S.* Typhimurium infection in mice

The oral infectious dose for Balb/C mice is some 10^5 *S.* Typhimurium [29]. After ingestion the bacteria reach the stomach, where a large majority is killed [30]. The few surviving bacteria pass from the stomach into the intestine, where they cross through the intestinal barrier into the lamina propria. Invasion is facilitated via endocytosis by microfold cells (M cells) in Peyer´s patches, or via sampling by dendritic cells directly from the gut lumen [31-35]. *Salmonella* causes systemic disease by infecting dendritic cells and CD18-expressing phagocytes, which disseminate the pathogen throughout the body. *S.* Typhimurium replicates then mainly in macrophages in mesenteric lymph nodes, spleen, liver and bone marrow [27, 36, 37]. *Salmonella* also colonizes the gall bladder epithelium. The bacteria are shedded through the bile duct into the intestine and are excreted with feces, thus completing the transmission cycle [38, 39].

After uptake of *Salmonella* by macrophages, the *Salmonella* containing phagosome initially follows normal phagosome maturation, but largely prevents fusion with lysosomes ([40-46], reviewed in [47]). Instead, the *Salmonella* containing vacuole (SCV) provides an intracellular replication niche for the pathogen, which is essential for systemic infections [48, 49]. Earlier studies suggested that the SCV forms an isolated compartment in the host cell, but recent results indicate active communication of the SCV with vesicles of host cell endo- and exocytosis [45, 50].

During the various steps in the infection cycle, *Salmonella* uses at least 200 different virulence genes. The chromosomal region called *Salmonella* pathogenicity island 1 (SPI1) encodes a type III secretion system (T3SS) and associated effectors that are secreted during infection [51]. SPI1 promotes gastroenteritis and induces phagocytosis in non-phagocytic cells through actin remodeling ([52], reviewed in [53]). For intracellular survival and growth, PhoPQ as well as virulence genes of the *Salmonella* pathogenicity island 2 (SPI2) are required [49, 54-56]. PhoPQ is a two-component system that can sense intracellular conditions [57, 58]. Upon uptake into the phagosome, it activates (amongst others) genes for resistance against oxidative stress and for modification of the outer membrane [59-61]. SPI2 encodes a second T3SS and associated

secreted effector proteins. The main function of SPI2 is modification of host vesicle trafficking to promote *Salmonella in vivo* growth ([62], reviewed in [63]).

1.5 Metabolism, virulence and *in silico* approaches

In addition to virulence genes, *Salmonella* survival and growth in host tissues depends on specific metabolic capabilities. Different sets of metabolic genes are expressed during gastroenteritis and systemic infections, with metabolic genes making up more than half of the proteins with detectable *in vivo* abundance [64, 65]. Defects in metabolism (anabolism/catabolism) reduce replication and virulence. For example, the inactivation of anabolic genes like *aroA*, *purA*, *asd, fabB*, *ribB* or catabolic genes of glycolysis (*pfkAB*) and citric acid cycle (*sucCD*, *mdh*) resulted in attenuated to avirulent *Salmonella* mutants ([48, 49, 64, 66-69], own unpublished data). The ability to use host metabolites as nutrient source or as electron acceptor is also a distinct competitive advantage or even a necessity for *Salmonella* virulence ([69-71] own unpublished data).

These examples demonstrate crucial importance of *Salmonella* metabolism for virulence. However, the various findings have not yet yielded an integrated quantitative understanding of metabolism during infection. Specifically, it is unclear:

- Which nutrients are available to *Salmonella* during infection, and what impact these nutrients have on *Salmonella in vivo* metabolism.
- How different metabolic pathways interact during infection.
- What the reasons for the remarkable robustness of *Salmonella* metabolism are [64].
- Whether the vast majority of metabolic genes with undetectable virulence contribution is really dispensable during infection [72, 73].

Resolving these issues requires a quantitative and comprehensive understanding of metabolism based on a combined experimental / *in silico* approach. To enable such an approach, all known metabolic reactions with associated metabolites and enzymes are listed in a computer-readable format (Systems biology markup language (SBML) [74]). The reactions are incorporated in the respective stoichiometries with charge and mass balance. Reaction irreversibility is also determined, based on reaction thermodynamics [75]. The result is a Biochemically, Genetically and Genomically structured genome-scale metabolic network reconstructions (BiGG database) (reviewed in [76]). A BiGG database is based on *in vitro* metabolism, which is well characterized in *Salmonella* ([77, 78], reviewed in [79]). Such an *in silico* database can be analyzed using a method called flux balance analysis (FBA). FBA predicts metabolic reactions that can yield biomass with all required components for cell growth in defined environmental conditions. The main advantage of FBA is that this method can be applied with high accuracy and predictive power to genome-scale metabolic networks [80-82]. FBA can also be used to predict the phenotypes of metabolic mutants, with recent *Escherichia coli* reconstructions predicting gene essentiality with an accuracy of over 90% ([80, 81, 83], own unpublished data). The utility of this approach is also emphasized by consistency of FBA predictions with proteome and transcriptome data [84].

As *Salmonella* metabolism and growth depends on available nutrients, this must be an integral part of any comprehensive model. However, data on relevant nutrients in infected host tissues remains fragmentary and quantitative data on nutrient supply are completely lacking. These knowledge gaps have so far severely limited applicability of *in silico* modeling of *Salmonella* and other pathogens.

1.6 Goal of the thesis

The goal of this thesis was to analyze *Salmonella in vivo* nutrition, genome scale metabolism and *in vivo* replication. To achieve this goal, four specific aims were followed:

1) To generate a reconstruction of *Salmonella* metabolism for enabling *in silico* analysis of *Salmonella* metabolism during infection (see 2.1).

2) To obtain quantitative nutrient availability data during infection and to incorporate these data in a comprehensive metabolism model (see 2.2).

3) To analyze extent and causes of apparent massive enzyme dispensability during infection (see 2.3).

4) To develop an unbiased method for large scale analysis of enzyme dispensability (see 2.4).

2. Results

2.1 A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella* Typhimurium LT2

BMC Syst Biol. 2011 Jan 18;5:8.

Thiele I, Hyduke DR, **Steeb B**, Fankam G, Allen DK, Bazzani S, Charusanti P, Chen FC, Fleming RM, Hsiung CA, De Keersmaecker SC, Liao YC, Marchal K, Mo ML, Özdemir E, Raghunathan A, Reed JL, Shin SI, Sigurbjörnsdóttir S, Steinmann J, Sudarsan S, Swainston N, Thijs IM, Zengler K, Palsson BO, Adkins JN, Bumann D

Summary:

Parallel efforts of separate groups can lead to different metabolic reconstructions (MRs) for the same organism. Differences can arise e.g. due to inconsistent literature interpretation or dissimilar reconstruction methods. The merging of different MRs leads to an improved consensus version, which can be used by the scientific community for more accurate *in silico* analysis of the organism in question.

In this paper, we presented a consensus version of a genome-scale metabolic network reconstruction of *Salmonella enterica* subspecies *enterica* serovar Typhimurium. In a community-driven effort of more than 20 experts in *S.* Typhimurium biology and systems biology, two independent metabolic reconstructions of *Salmonella* metabolism (BRecon and AJRecon) were merged to obtain a consensus reconstruction (STMv1.0). Both MRs were based on related *E. coli* reconstructions from the same research group, thus comparison was facilitated by similar syntax, metabolite and reaction labeling. Identical reactions were directly included into the consensus reconstruction, whereas dissimilar reactions were included or excluded after careful curation. The resulting consensus reconstruction STMv1.0 (i) possessed three compartments (cytosol, periplasm, exctracellular space), (ii) included reaction directionality based on thermodynamic calculations, (iii) can be used as a mathematical model for flux balance analysis (FBA) and (iv) predicted growth capability on diverse compounds with an overall accuracy of 88%.

Statement of my work:

I generated one of the two underlying reconstructions of *Salmonella* metabolism (BRecon), based on the *E. coli* reconstruction iAF1260 [81]. BRecon was merged with the reconstruction AJRecon to obtain the here presented consensus reconstruction STMv1.0. I participated in all phases of the generation of this consensus model (preparation phase, jamboree in Reykjavik (Iceland), literature curation and reconstruction finalization).

BMC Systems Biology

RESEARCH ARTICLE Open Access

A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella* Typhimurium LT2

Ines Thiele[1,2†], Daniel R Hyduke[3†], Benjamin Steeb[4], Guy Fankam[3], Douglas K Allen[5], Susanna Bazzani[6], Pep Charusanti[3], Feng-Chi Chen[7], Ronan MT Fleming[1,8], Chao A Hsiung[7], Sigrid CJ De Keersmaecker[9], Yu-Chieh Liao[7], Kathleen Marchal[9], Monica L Mo[3], Emre Özdemir[10], Anu Raghunathan[11], Jennifer L Reed[12], Sook-Il Shin[11], Sara Sigurbjörnsdóttir[13], Jonas Steinmann[13], Suresh Sudarsan[14], Neil Swainston[15,16], Inge M Thijs[9], Karsten Zengler[3], Bernhard O Palsson[3], Joshua N Adkins[17], Dirk Bumann[4*]

Abstract

Background: Metabolic reconstructions (MRs) are common denominators in systems biology and represent biochemical, genetic, and genomic (BiGG) knowledge-bases for target organisms by capturing currently available information in a consistent, structured manner. *Salmonella enterica* subspecies I serovar Typhimurium is a human pathogen, causes various diseases and its increasing antibiotic resistance poses a public health problem.

Results: Here, we describe a community-driven effort, in which more than 20 experts in *S.* Typhimurium biology and systems biology collaborated to reconcile and expand the *S.* Typhimurium BiGG knowledge-base. The consensus MR was obtained starting from two independently developed MRs for *S.* Typhimurium. Key results of this reconstruction jamboree include i) development and implementation of a community-based workflow for MR annotation and reconciliation; ii) incorporation of thermodynamic information; and iii) use of the consensus MR to identify potential multi-target drug therapy approaches.

Conclusion: Taken together, with the growing number of parallel MRs a structured, community-driven approach will be necessary to maximize quality while increasing adoption of MRs in experimental design and interpretation.

Background

The evolution of antibiotic resistance by a variety of human pathogens is a looming public health threat [1,2]. *Salmonella* is a major human pathogen and a model organism for bacterial pathogenesis research [3]. *S. enterica* subspecies I serovar Typhimurium (*S.* Typhimurium) is the principle subspecies employed in molecular biology and its variants are causative agents in gastroenteritis in humans. The publication of the annotated genome for *S.* Typhimurium LT2 provided a foundation for numerous applications, such as drug discovery [4]. Previous efforts to systematically identify candidate drug targets within metabolism did not result in a plethora of new candidates, due to the robustness and redundancy of *S.* Typhimurium's metabolic network [5]. Since new single protein targets are missing, we need to target multiple proteins conjointly. Unfortunately, antibiotic regimens, which require multiple targets to be hit simultaneously, have an increased probability of the pathogen evolving resistance relative to a single target therapy. However, the continuous clinical success of the combination of beta-lactams and beta-lactamase inhibitors actually demonstrates that inhibitor combinations can be successful even if each individual inhibitor is non-effective on its own. The robustness inherent to *S.* Typhimurium's metabolic network imposes combinatorial challenges for *in vitro* and *in vivo* approaches to identify synthetic lethal genes sets (*i.e.*, experimental enumeration of all synthetic lethal pairs in *S.* Typhimurium would require the creation of

* Correspondence: dirk.bumann@unibas.ch
† Contributed equally
[4]Infection Biology, Biozentrum, University of Basel, Basel, Switzerland
Full list of author information is available at the end of the article

~500,000 double gene deletion strains, see below). Employing a systems biology network perspective could facilitate their identification.

GEnome scale Network REconstructions (GENRE) [6] represent biochemical, genetic, and genomic (BiGG) knowledge-bases [7] for target organisms; and have been developed for expression [8,9], metabolic [6,10], regulatory [11], and signaling [12,13] networks. Metabolic reconstructions (MRs) are the most developed out of the four GENRES. The metabolic network reconstruction process is well established [14] and has been used for various biotechnological and biomedical applications [15,16]. Given the rapidly growing interest in MRs and modeling, parallel reconstruction efforts for the same target organism have arisen and resulted in alternative MRs for a number of organisms [17-23]. These parallel MRs may vary in content and format due to differences in reconstruction approaches, literature interpretation, and domain expertise of the reconstructing group. Subsequent network comparison and discoveries are hampered by these differences. Consequently, the need for a community approach to divide the substantial effort required in reconciling and expanding these MRs has been formulated [17].

Results and Discussion

Salmonella, a reconstruction jamboree for an infectious disease agent

In June 2008, it became apparent that two MRs were being assembled by two different research groups [20] (Bumann, unpublished data). Subsequently, a *Salmonella* reconstruction jamboree was held at the University of Iceland, Reykjavik, from September 5th to 6th, 2008. The jamboree team consisted of over 20 experts in microbiology, proteomics, *Salmonella* physiology, and computational modeling. Based on the experience with the yeast reconstruction jamboree [17], a methodology was devised to increase the efficiency of community-based network reconstruction [24] and applied to the *Salmonella* reconstruction jamboree.

The goal of a network reconstruction jamboree is to provide a 2-D genome annotation that is of higher quality than it may be achieved by bioinformatic analyses alone [24,25]. The objective of this jamboree was to re-evaluate, reconcile, and expand the currently available MRs for *S*. Typhimurium with a focus on virulence. Furthermore, we aimed to include standard identifiers for reconstruction metabolites, reactions, and genes to facilitate subsequent mapping of 'omics' data. The starting MRs were AJRecon (a variant is published in [20]) and BRecon (D. Bumann, unpublished data), which were derived from published *E. coli* MRs, iJR904 [26] and iAF1260 [27], respectively, and their contents were modified to account for *Salmonella*-specific properties; i.e., transport and enzymatic reactions not present in *Salmonella* were removed and the proteins associated with the reactions were modified to contain proteins present in *S*. Typhimurium LT2.

Comparison of two metabolic reconstructions for S. Typhimurium

We developed an automatic approach to initiate the reconciliation of the two MRs by converting their metabolites and reactions into a common language (Figure 1). The MR contents were grouped into three categories: (1) identical, (2) similar, and (3) dissimilar. A similar reaction was one, in which there was a minor discrepancy, such as reaction reversibility, a missing reactant or product, or a difference in associated enzyme(s). Dissimilar reactions were those with distinct sets of reactants and products, and often represented metabolic reactions that were not included in one of the starting MRS. The identical content was transferred to the consensus MR without further evaluation. The similar and dissimilar content was evaluated at the jamboree. Genes and proteins associated with the reactions were also carefully compared and refined where necessary. At its end, the meeting yielded an approximately 80% reconciled consensus reconstruction. The remaining discrepancies were manually curated by the Bumann and Palsson groups following the jamboree meeting.

Initial comparison revealed that there were 760 reactions common to the starting MRs while 521 and 1684 reactions were unique to AJRecon and BRecon, respectively (Additional file 1 Table S1). Some of these differences could be explained by changes introduced to *the E. coli* MR when it was converted from its earlier version, iJR904 [26], to the most recent version, iAF1260 [27] (i.e., explicit definition of a periplasm compartment; more detailed fatty acid metabolism).

Characteristics of the Salmonella Consensus Reconstruction

The resulting knowledge-base, STM_v1.0 (Table 1; Additional file 2; Additional file 1 Table S2), represents the final product of a community-effort to develop a detailed MR of *S*. Typhimurium. STM_v1.0 integrates the novel and common features of the starting MRs into a vetted, well-documented consensus knowledge-base, capturing currently available BiGG knowledge about *S*. Typhimurium. Key features of STM_v1.0 include i) accounting for the periplasm as a compartment between the extracellular space and cytoplasm; ii) *Salmonella*-specific virulence characteristics, such as iron chelation by salmochelin and serovar Typhimurium LT2 O-antigen production; iii) the possibility to employ the consensus MR as mathematical, predictive model; and iv) comprehensive support data for reactions and associated genes (Additional file 1 Table S2a). Some

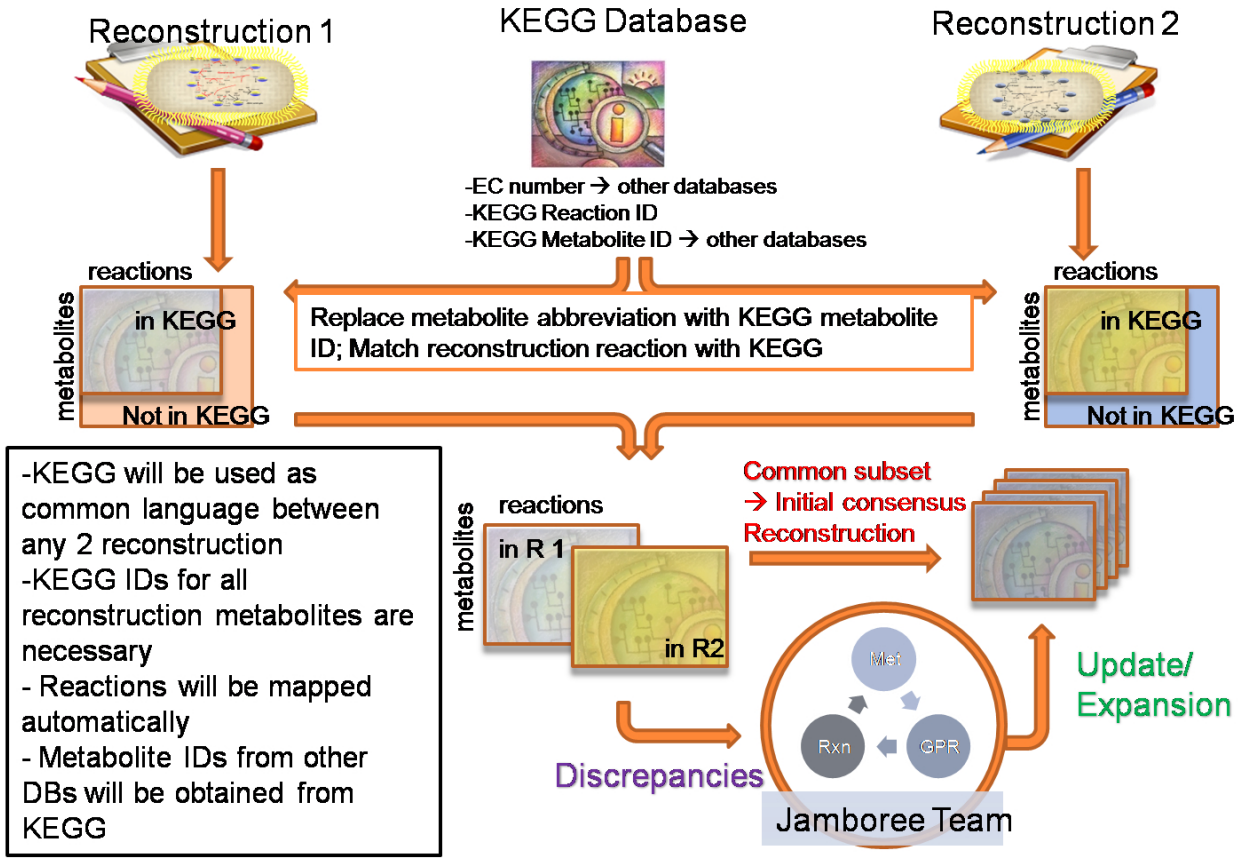


Figure 1 Approach to reconcile two metabolic reconstructions (MR). This figure illustrates the automated comparison tool developed for the Salmonella reconstruction jamboree. Both MRs are translated into a common language (based on KEGG [44]). Metabolites and reactions that could not be mapped onto KEGG were subject to manual evaluation by the jamboree team. The overlapping part of the MRs was directly moved into the consensus MR while reactions and metabolites unique to a MR were evaluated manually. This approach can be readily applied to comparison of any two MRs.

information was excluded from STM_v1.0, such as the 26 dipeptide and tripeptide transport/digestion reactions that are present in AJRecon, as they represent generic compounds. Accounting for all potential consumable oligopeptides would make computational analysis intractable or unnecessarily difficult. Appropriate oligopeptides may be manually added to STM_v1.0 to represent a specific growth environment. We also attempted to exclude reactions that were included to fit some growth data [28], but were contrary to other observations [20,29] as was the case for growth with D-aspartic acid [30] as the sole carbon source which

Table 1 Basic Statistics for the original and the consensus reconstructions

	AJRecon [20]	BRecon	iMA945* [21]	Consensus (new data)
Genes	1,119	1,222	945	1,270
Network reactions	1,079	2,108	1,964	2,201
-Transport reactions	200	575	726	738
Biochemical reactions	879	1,533	1,238	1,463
Metabolites (unique)	754	1,084	1,035	1,119
Compartments	Cytosol, extracellular space	Cytosol, periplasm, extracellular space	Cytosol, periplasm, extracellular space	Cytosol, periplasm, extracellular space

*Not included in consensus reconstruction. See text for details.

requires an unknown transporter and an unknown aspartate racemase [31].

Additionally, we evaluated the reaction directionality of consensus MR reactions by considering thermodynamic properties of participating metabolites. In the case that a thermodynamic prediction was inconsistent with experimental evidence, the experimental evidence was followed. Thermodynamic predictions are made using the knowledge that is available [45], and incorrect predictions highlight gaps in our knowledge of biology.

A bacterial MR often includes a biomass reaction that lists all known biomass precursors and their fractional contribution necessary to produce a new bacterial cell in a given environment. The individual biomass constituents of a *S.* Typhimurium cell have been measured [20], and adapted for the consensus reconstruction by accounting for the changes in naming and compartments introduced during reconciliation (Additional file 1 Table S3c).

Comparison with a third metabolic reconstruction of S. Typhimurium

After finishing the consensus reconstruction, a third metabolic reconstruction (iMA945) was published [21]. Similar to one of our starting MRs (BRecon), iMA945 was built by using homology, and other bioinformatics criteria [32], starting from the *E. coli* metabolic reconstruction (iAF1260). Gaps in iMA945 were detected and filled with GapFind and GapFill, respectively [33]; and iMA945's content was further augmented by the GrowMatch algorithm [34] to fit experimental measurements. These automated optimization methods are excellent tools for identifying gaps in network reconstructions and proposing candidate reactions to fill these gaps and fit the model to growth data, however, they often do not associate genes with the candidate reactions. The candidate reactions are typically taken from a universal reaction database (such as KEGG) that includes pathways from all domains of life, thus candidate reactions proposed by these methods should be taken as hypotheses and require additional validation from published literature or direct experimental evidence.

We performed a preliminary comparison between STM_v1.0 and iMA945. However, we did not reconcile iMA945 with the consensus reconstruction, as this will require detailed evaluation of the discrepancies in a subsequent jamboree meeting. Overall, 2,057 reactions were present in both the consensus reconstruction and iMA945, of which 1,706 reactions have identical gene-protein-reaction (GPR) associations (Additional file 1 Table S2d). A total of 26 reactions had identical reaction identifiers but different reactions (*e.g.*, different reactants, products, stoichiometry, or directionality: reversible, forward only, backward only) and GPR associations. There were a total of 629 distinct reaction ids between STM_v1.0 and iMA945: 446 were unique to STM_v1.0 and 183 to iMA945. Of the 183 reactions flagged as unique to iMA945, the majority represents reactions that were intentionally excluded from the consensus reaction (*e.g.*, 45 dipeptide exchange, transport, and peptidase reactions and >60 additional exchange, transport, and enzymatic reactions not supported by literature). Some of the distinct reactions, such as adenosylcobalamin phosphate synthase, were due to different metabolite and reaction identifiers. No bibliomic data were included in iMA945, so it was not possible to assess whether the reactions were inserted by the automated gap-filling methods or supported by additional evidence. The 446 reactions unique to STM_v1.0 include *Salmonella*-specific chelators, O-antigens, and lipid modifications that were not present in the starting network derived from the *E. coli* MR (iAF1260). Overall, the core metabolic network is similar between STM_v1.0 and iMA945, which is expected as the draft scaffolds for both MRs were derived from *E. coli* MRs and *S.* Typhimurium has a notable metabolic homology with *E. coli*; however, STM_v1.0 includes over 300 more genes than iMA945 and includes a variety of *Salmonella*-specific reactions that are essential for virulence and could serve as coupling points for constructing a host-pathogen model.

Metabolic Network Reconstruction Assessment

To assess the utility of a mathematical approximation of reality, it is essential to determine the consistency of the model's predictions with real-world benchmarks. In the case of MRs, comparing experimental growth data with predicted biomass production is a commonly employed metric in benchmarking metabolic models [14]. Although biomass production is a commonly employed metric, the results should always be taken with a grain of salt; for instance, it is possible to improve the fitting of a model's predictions to growth data by including enzymatic reactions for which no evidence exists or which are contrary to published experimental observations. The reconstruction committee chose not to include invalidated enzymatic reactions that improved the fit between growth predictions and experimental observations; the failings of the model's predictions highlight areas where knowledge is lacking and experimental undertakings could identify new knowledge.

For *S.* Typhimurium, there is a wealth of experimental growth data [29]. Overall, we found good agreement between the qualitative growth phenotype predictions and the experimental data (Table 2 Additional file 1 Table S4); with the notable exception of sulfur metabolism where the prediction accuracy was about 40%. As we are becoming increasingly aware of the importance of sulfur-related metabolism in host-pathogen

Table 2 Growth benchmark results

Source (accuracy)	Prediction	Experiment	
		Growth	No Growth
Carbon	Growth	79	9
(118/133)	No Growth	6	39
Nitrogen	Growth	28	5
(57/64)	No Growth	2	29
Phosphate	Growth	24	0
(24/25)	No Growth	1	0
Sulfur	Growth	6	0
(8/14)	No Growth	6	2

interactions [35-38], the deficiencies in our knowledge highlighted by this analysis represent viable targets for experimental enquiry. For the carbon and nitrogen sources accessible by AJRecon our results were comparable, however STM_v1.0 has the ability to metabolize 20 carbon sources and 15 nitrogen sources not accessible to AJRecon. The additional metabolic capabilities of STM_v1.0 are due, in part, to the presence of ~200 additional gene products in STM_v1.0.

Gene Essentiality Simulations

To combat the rise in antibiotic-resistant pathogens, it is crucial to identify new drug targets. Genes or sets of genes that are essential for growth are potential drug targets. To identify novel drug targets in STM_v1.0, we performed single and double gene deletion studies. We identified 201 essential genes in M9/glc, 144 of which were also essential in LB (Additional file 1 Table S5a). The synthetic lethal gene pair simulations were performed using only genes that were found to be non-essential in the condition of interest (Additional file 1 Table S6). In M9/glc, there were 87 synthetic lethal gene-pairs comprised of 102 unique genes. For *E. coli*, Suthers *et al.* [39] predicted 86 synthetic lethal gene-pairs, however, there were only 83 unique genes involved. In LB, there were 56 synthetic lethal gene-pairs comprised of 76 unique genes. Interestingly, 10 of LB synthetic lethal genes were also essential in M9/glc and were members of 12 of the LB synthetic lethal gene-pairs. The very small fraction of essential synthetic lethal gene pairs (< 100 synthetic lethalities out of >500,000 possibilities - assuming approx. 1000 non-essential metabolic genes) emphasizes the robustness of *S*. Typhimurium's metabolic network, which has previously been noted [5].

Candidate drug targets

Our observed, very small number of synthetic lethal pairs in STM_v1.0 indicates that antimicrobial regimens may need to target more than two elements to be effective. Unfortunately, it will take less time for a pathogen to evolve a solution to a conjoint two-target antimicrobial strategy compared to a single-target strategy. To reduce the probability of a pathogen evolving resistance to a conjoint two-target strategy, it may be plausible to employ a combination of two-target strategies. Although a combination approach may be suitable for dealing with antibiotic resistance, there are potential shortcomings associated with clearance and toxicity because all the components of a regimen must reach a target at a specific time with the requisite concentrations. Despite these difficulties, multi-component, multi-target drugs are becoming standard therapeutics for complex diseases, including cancer, diabetes, and infectious diseases [40]. Experimental identification and characterization of therapeutic strategies that require multiple targets for effectiveness is a resource intensive undertaking (*e.g.*, creating over 500,000 double mutant strains). An *in silico* approach using an MR, such as STM_v1.0, could be implemented to prioritize the experiments by indicating which multi-target therapies would adversely affect the pathogen's metabolic capabilities.

As mentioned above, the synthetic gene deletion analysis yielded 56 synthetic lethal gene pairs disrupting growth of *S*. Typhimurium *in silico*. We grouped these gene pairs based on different criteria to assess their potential value as multi-drug targets (Figure 2). It is notable that five gene pairs are between protein complexes while a further three gene pairs are between genes involved in the same pathway - this indicates the presence of a layer of 'redundancy' for the enzyme or pathway that confers protection against a single-target therapy. Moreover, three of the genes involved in gene pairs are known to be essential for virulence, but not for growth, and have known inhibitors based on BRENDA [41]. This structured overview of *in silico* synthetic lethal gene pairs identified numerous candidate drug targets many of which have known inhibitors. In subsequent studies, these model-generated hypotheses need to be tested and validated.

Additional gene products shown to play a central role in virulence yet are not essential for growth in laboratory conditions or do not have an unequivocal functional annotation represent additional therapeutic targets. These gene products could serve as potential points for manipulating host metabolism [38], could be essential for metabolism in the host environment (*e.g., Salmonella*-containing vacuoles are nutrient poor) [42], and will represent an energy and materials demand when creating integrated metabolic and expression reconstructions [8,9]. Recent examples of relevant gene products that have not been annotated but are crucial for virulence include gene products STM3117-STM3120 [43]; as the

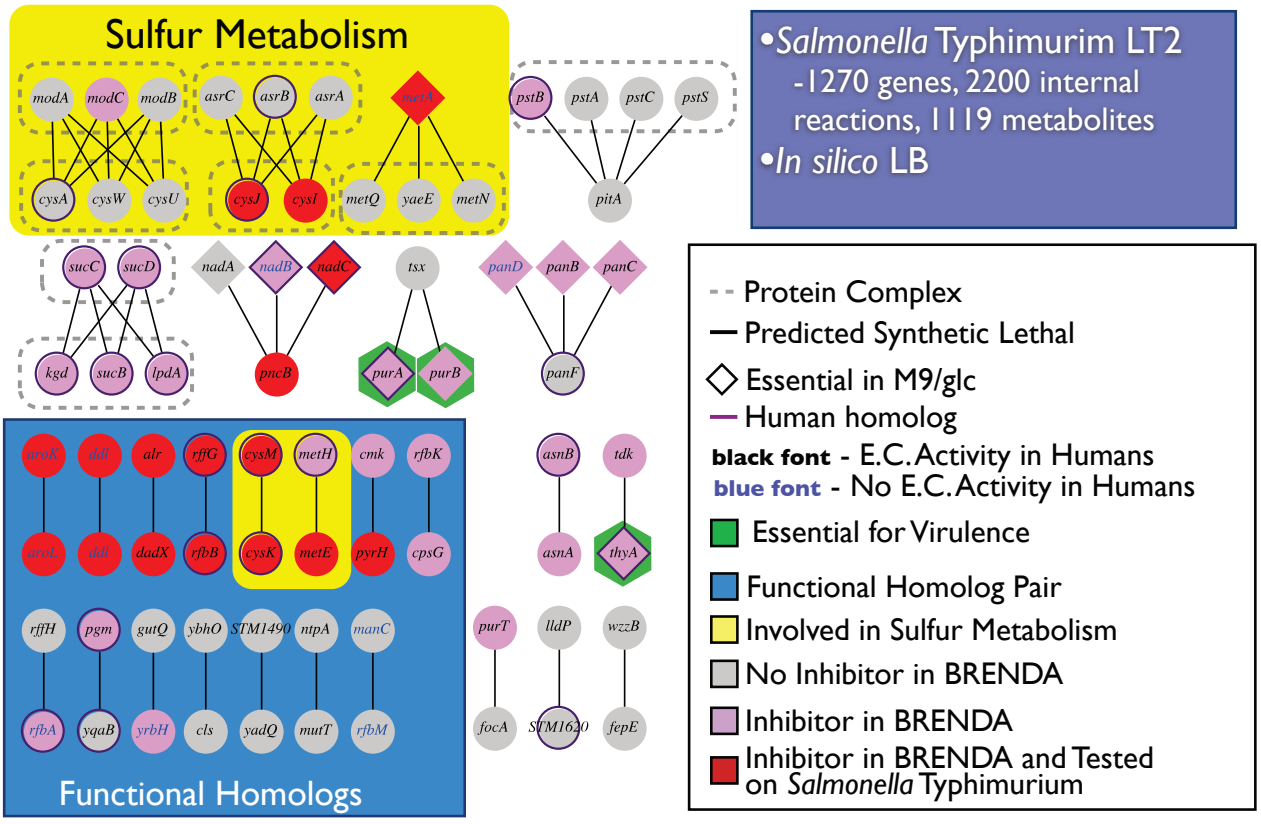


Figure 2 Candidate drug targets. The figure contains all predicted synthetic lethal interactions for STM_v1.0 in LB medium. A line connecting two genes represents a synthetic lethal pair. A group of genes surrounded by a dashed box represents a protein complex requiring all enclosed elements to function. Yellow background means associated with sulfur metabolism. Blue background indicates that the synthetic lethal pairs are functional homolog's. Red gene means that there is a chemical known to inhibit the gene-product in STM_v1.0.

metabolic functions of unannotated genes are elucidated, they will need to be incorporated into future revisions of the MR.

Conclusions

Taken together, the community-developed consensus MR is a curated reconstruction with the combined properties of the starting MRs and new information that was added during and after the reconstruction jamboree. The expanded metabolic versatility with a focus on virulence, updated annotation, including corrections, and curation of hundreds of additional reactions, genes, and metabolites by a community of experts present in STM_v1.0 highlights the value of a community-based approach. Another MR for *S.* Typhimurium was published after the jamboree [21], which was also based on an *E. coli* MR [27]. The reconciliation with this third MR will need to be done in subsequent jamboree meetings, which will also lead to a further expansion of knowledge and data included in the consensus knowledge-base. The publication of the third MR for *S.* Typhimurium emphasizes the importance and the value of the effort presented in this report as well as the need for additional outreach when assembling jamboree committees.

Methods

Metabolic network reconstructions of Salmonella enterica serovar *Typhimurium LT2*

The starting reconstructions, AJRecon and BRecon, were built on scaffolds derived from published *E. coli* MRS. AJRecon is a pre-publication version of iRR1083 [20], and was based on iJR904 [26]. For its scaffold, BRecon (Bumann, unpublished) employed iAF1260 [27]- a direct descendent of iJR904. The two reconstructions, differ in content due to: (1) different components being targeted for manual curation (e.g., BRecon extended Fe chelation and AJRecon extended lipid production), and (2) differences in *E. coli* MRs that were used as comparative genomics scaffolds for initializing the *Salmonella* MRs (e.g., iAF1260 accounted for the periplasm whereas its ancestor did not).

Method for community-based network reconstruction

There are three essential phases for community-based MR development: (1) preparation, (2) jamboree, and

(3) reconstruction finalization [24]. The preparation and finalization phases are carried out by a small contingent of researchers, whereas, the collective knowledge of the community is harnessed during the jamboree. In the preparation phase, the two MRs were compared as described below in terms of metabolites, reactions, and gene-protein-reaction associations (GPRs). Overlapping content between both original MRs was directly moved into the consensus MR (Additional file 1 Table S1). Discrepancies in the listed three areas were presented to the jamboree team, which was split into three groups: metabolite curation, reaction curation group, and GPR curation group. The metabolites group curated the list of all metabolites present in either original MR for i) protonation state of metabolites at physiological pH, ii) missing metabolite identifiers: KEGGID, PubChemID, ChEBI ID, and iii) comparison of neutral formulae in reconstruction and metabolite databases. The reaction group was responsible for identifying evidence for orphan reactions in either original MR with and without a KEGG reaction ID. Reactions without a KEGG ID had to be extensively audited as there were no database evidences for the correctness of the reaction mechanisms. The GPR group had to resolve the discrepancies in GPR assignments using genome databases and literature. Each team evaluates their problem set based on evidence within the consensus MR and available resources (literature, databases, and annotations). Items that are not adequately addressed during the jamboree are subject to extensive manual curation during the MR finalization phase. The finalization phase includes: (1) manual curation, (2) benchmarking the consensus MR against experimentally-derived phenotypic data, and (3) MR dissemination. The consensus MR is expected to be maintained, updated and expanded in subsequent reconstruction jamborees.

Metabolic Reconstruction Reconciliation

Reconciling multiple MRs requires that the MRs' contents employ a common nomenclature so that the contents may be compared. For this work, we employed the KEGG database [44] as the source of common identifiers (Figure 1); although all of the reactions and metabolites in KEGG may not be accurate or complete, KEGG has the benefit of being an extensive, freely accessible resource used by the broader biological community. The complete consensus reconstruction can be found in Additional file 1 Table S6 and in Additional file 2 as an SBML file.

Thermodynamic directionality

Thermodynamic directionality for each reaction was calculated as described in [45]. Briefly, assuming a temperature of at 310.15 K, intracellular pH of 7.7, extracellular/periplasmic pH of 7.0, and a concentration range of 0.01-20 mM, we calculated upper and lower bounds on transformed reaction Gibbs energy, and assigned reaction directionality accordingly. Transport reactions were not subject to thermodynamic consistency analysis as there is still uncertainty associated with the directionality prediction of transmembrane transport.

Conversion of reconstruction into a mathematical model

The conversion of a reconstruction into a mathematical model has been described in detail elsewhere [14]. The unit of reaction fluxes was defined as mmol/g_{DW}/hr.

Phenotypic assessment

Flux balance analysis [46] was employed to assess the STM_v1.0 model's ability to correctly predict biomass production in a variety of limiting conditions. The accuracy of the model was assessed by comparing the predictions to benchmarks drawn from experimental data [20,29]. In this assessment, there are four possible observations: (1) STM_v1.0 model correctly predicts growth (G/G), (2) STM_v1.0 model incorrectly predicts growth (G/NG), (3) STM_v1.0 model correctly predicts no growth (NG/NG), and (4) STM_v1.0 model incorrectly predicts no growth (NG/G). For a prediction to be counted as a true positive (G/G) or true negative (NG/NG), the prediction needed to match one or more experimental observations. The predictions were first compared with the Biolog phenotype microarray (PM) data http://www.biolog.com. False positive predictions (G/NG) and false negative predictions (NG/G) were then compared with the data from Gutnick *et al.* [29] and references cited in Ragunathan *et al.* [20]. For limiting conditions not represented in the PM, predictions were only compared with data from Gutnick *et al.* [29] or cited in Ragunathan *et al.* [20].

Gene essentiality analysis

The gene deletion studies were performed by converting STM_v1.0 into a stoichiometric model and performing flux balance analysis [46]. For each gene, or gene pair, the associated reaction(s) were disabled ($v_{min,\ i} = v_{max,\ i} = 0$ mmol.gDW^{-1}.hr^{-1}) and the ability of the model to produce biomass was assessed, i.e., the biomass reaction was chosen as the objective function and maximized.

All simulations were performed using the COBRA Toolbox v2.0 [47] using Matlab (Mathworks, Inc) as the programming environment, and Tomlab (TomOpt, Inc) as the linear programming solver.

Additional material

Additional file 1: Consensus MR. This xlsx file contains the consensus reconstruction and simulation setup/results. - Table S1. Statistics for automated reconciliation of starting reconstructions. - Table S2.

Consensus Reconstruction in SBML format. - Table S3a. M9/glc. - Table S3b. LB. - Table S3c. Biomass. - Table S4. Growth benchmark errors. - Table S5a. All Lethal deletion predictions. - Table S5b. Single Deletion/ Virulence. - Table S6a. LB Synthetic Lethal. - Table S6b. M9 Synthetic Lethal.

Additional file 2: Consensus MR in SBML format. Consensus MR as a computational model in SBML format.

Acknowledgements
The authors would like to thank R. Archila and K.C. Soh for participation at the opening day of the reconstruction jamboree. I.T. would like to thank M. Herrgard for the valuable discussions. This work was supported in part by the National Institute of Allergy and Infectious Diseases NIH/DHHS through interagency agreement Y1-AI-8401-01. I.T. was supported in part by a Marie Curie International Reintegration Grant within the 7th European Community Framework Program (PIRG05-GA-2009-249261).

Author details
[1]Center for Systems Biology, University of Iceland, Reykjavik, Iceland. [2]Faculty of Industrial Engineering, Mechanical Engineering & Computer Science University of Iceland, Reykjavik, Iceland. [3]Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. [4]Infection Biology, Biozentrum, University of Basel, Basel, Switzerland. [5]USDA-ARS, Plant Genetics Research Unit, Donald Danforth Plant Science Center, St Louis, MO, USA. [6]Technical University Braunschweig, Institute for Bioinformatics & Biochemistry, Braunschweig, Germany. [7]Division of Biostatistics and Bioinformatics, Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan. [8]Science Institute, University of Iceland, Reykjavik, Iceland. [9]Centre of Microbial and Plant Genetics, Department of Microbial & Molecular Systems, Katholieke Universiteit Leuven, Leuven, Belgium. [10]Laboratory of Computational Systems Biotechnology, Ecole Polytechnique Fédérale de Lausanne, Swiss Institute of Bioinformatics, Lausanne, Switzerland. [11]Department of Infectious Diseases, Mount Sinai School of Medicine, New York City, NY, USA. [12]Department of Chemical & Biological Engineering, University of Wisconsin-Madison, Madison, WI, USA. [13]Faculty of Life & Environmental Sciences, University of Iceland, Reykjavik, Iceland. [14]Department of Biochemical and Chemical Engineering, Technische Universität Dortmund, Dortmund, Germany. [15]School of Computer Science, The University of Manchester, Manchester, UK. [16]The Manchester Centre for Integrative Systems Biology, Manchester Interdisciplinary Biocentre, The University of Manchester, Manchester, UK. [17]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA.

Authors' contributions
IT, DRH, BOP, JNA, and DB conceived the study. BS and DB compiled the BRecon. IT and DRH compiled the consensus MR. IT, DRH, BOP, and DB wrote the manuscript. GF and IT designed and performed initial MR comparisons. RMTF and DRH performed thermodynamic directionality analysis. DHR and IT carried out the computational analysis of the consensus MR. IT, BOP, DB, BS, DKA, SB, PC, FCC, RMTF, CAH, SCJK, YCL, KM, MLM, EÖ, AR, JLR, SIS, SS, JS, SS, NS, IMT, KZ, BOP, JNA, DB actively participated during and/or after the metabolic reconstruction jamboree to generate content for the consensus MR. All authors read and approved the final manuscript.

Received: 26 May 2010 Accepted: 18 January 2011
Published: 18 January 2011

References

1. Bjorkman J, Hughes D, Andersson DI: **Virulence of antibiotic-resistant Salmonella typhimurium.** *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95(7)**:3949-3953.
2. Norrby SR, Nord CE, Finch R: **Lack of development of new antimicrobial drugs: a potential serious threat to public health.** *The Lancet infectious diseases* 2005, **5(2)**:115-119.
3. Ohl ME, Miller SI: **Salmonella: a model for bacterial pathogenesis.** *Annual review of medicine* 2001, **52**:259-274.
4. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK: **Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.** *Nature* 2001, **413(6858)**:852-856.
5. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, Mann M, Bumann D: **Robust Salmonella metabolism limits possibilities for new antimicrobials.** *Nature* 2006, **440(7082)**:303-307.
6. Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO: **Reconstruction of biochemical networks in microorganismS.** *Nature reviews* 2009, **7(2)**:129-143.
7. Palsson BO: **Systems biology: properties of reconstructed networks.** New York: Cambridge University Press; 2006.
8. Thiele I, Jamshidi N, Fleming RM, Palsson BO: **Genome-scale reconstruction of Escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization.** *PLoS Comput Biol* 2009, **5(3)**:e1000312.
9. Thiele I, Fleming RM, Bordbar A, Schellenberger J, Palsson BO: **Functional characterization of alternate optimal solutions of Escherichia coli's transcriptional and translational machinery.** *Biophysical journal* 2010, **98(10)**:2072-2081.
10. Reed JL, Famili I, Thiele I, Palsson BO: **Towards multidimensional genome annotation.** *Nature reviews* 2006, **7(2)**:130-141.
11. Gianchandani EP, Joyce AR, Palsson BO, Papin JA: **Functional States of the genome-scale Escherichia coli transcriptional regulatory system.** *PLoS Comput Biol* 2009, **5(6)**:e1000403.
12. Hyduke DR, Palsson BO: **Towards genome-scale signalling-network reconstructions.** *Nature reviews* 2010, **11(4)**:297-307.
13. Li F, Thiele I, Jamshidi N, Palsson BO: **Identification of potential pathway mediation targets in Toll-like receptor signaling.** *PLoS Comput Biol* 2009, **5(2)**:e1000292.
14. Thiele I, Palsson BO: **A protocol for generating a high-quality genome-scale metabolic reconstruction.** *Nature protocols* 2010, **5(1)**:93-121.
15. Feist AM, Palsson BO: **The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*.** *Nat Biotech* 2008, **26(6)**:659-667.
16. Oberhardt MA, Palsson BO, Papin JA: **Applications of genome-scale metabolic reconstructionS.** *Molecular systems biology* 2009, **5**:320.
17. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Bluthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novère N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasić I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kirdar B, Penttilä M, Klipp E, Palsson BØ, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology.** *Nat Biotechnol* 2008, **26(10)**:1155-1160.
18. Jamshidi N, Palsson BO: **Investigating the metabolic capabilities of Mycobacterium tuberculosis H37Rv using the in silico strain iNJ661 and proposing alternative drug targetS.** *BMC systems biology* 2007, **1**:26.
19. Beste DJ, Hooper T, Stewart G, Bonde B, Avignone-Rossa C, Bushell ME, Wheeler P, Klamt S, Kierzek AM, McFadden J: **GSMN-TB: a web-based genome-scale network model of Mycobacterium tuberculosis metabolism.** *Genome biology* 2007, **8(5)**:R89.
20. Raghunathan A, Reed J, Shin S, Palsson B, Daefler S: **Constraint-based analysis of metabolic capacity of Salmonella typhimurium during host-pathogen interaction.** *BMC systems biology* 2009, **3**:38.
21. Abuoun M, Suthers PF, Jones GI, Carter BR, Saunders MP, Maranas CD, Woodward MJ, Anjun MF: **Genome scale reconstruction of a Salmonella metabolic model: comparison of similarity and differences with a commensal Escherichia coli strain.** *J Biol Chem* 2009, **284(43)**:29480-8.
22. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BO: **Global reconstruction of the human metabolic network based on genomic and bibliomic data.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104(6)**:1777-1782.
23. Ma H, Sorokin A, Mazein A, Selkov A, Selkov E, Demin O, Goryanin I: **The Edinburgh human metabolic network reconstruction and its functional analysis.** *Molecular systems biology* 2007, **3**:135.
24. Thiele I, Palsson BO: **Reconstruction annotation jamborees: a community approach to systems biology.** *Molecular systems biology* 2010, **6**:361.
25. Palsson BO: **Two-dimensional annotation of genomes.** *Nat Biotechnol* 2004, **22(10)**:1218-1219.

26. Reed JL, Vo TD, Schilling CH, Palsson BO: **An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR).** *Genome biology* 2003, **4(9)**:R54.51-R54.12.
27. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO: **A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information.** *Molecular systems biology* 2007, **3**:121.
28. **Biolog Inc.** [http://www.biolog.com/].
29. Gutnick D, Calvo JM, Klopotowski T, Ames BN: **Compounds which serve as the sole source of carbon or nitrogen for Salmonella typhimurium LT-2.** *Journal of bacteriology* 1969, **100(1)**:215-219.
30. Bayne HG, Stokes JL: **Amino acid metabolism of Salmonellae.** *Journal of bacteriology* 1961, **81**:126-129.
31. Oikawa T: **Alanine, Aspartate, and Asparagine Metabolism in Microorganisms.** In *Amino Acid Biosynthesis ~ Pathways, Regulation and Metabolic Engineering. Volume 5.* Edited by: Wendisch V. Berlin/Heidelberg: Springer; 2007:273-288.
32. Suthers PF, Dasika MS, Kumar VS, Denisov G, Glass JI, Maranas CD: **A genome-scale metabolic reconstruction of Mycoplasma genitalium, iPS189.** *PLoS Comput Biol* 2009, **5(2)**:e1000285.
33. Satish Kumar V, Dasika MS, Maranas CD: **Optimization based automated curation of metabolic reconstructions.** *BMC Bioinformatics* 2007, **8**:212.
34. Kumar VS, Maranas CD: **GrowMatch: an automated method for reconciling in silico/in vivo growth predictions.** *PLoS Comput Biol* 2009, **5(3)**:e1000308.
35. Justino MC, Almeida CC, Goncalves VL, Teixeira M, Saraiva LM: **Escherichia coli YtfE is a di-iron protein with an important function in assembly of iron-sulphur clusters.** *FEMS microbiology letters* 2006, **257(2)**:278-284.
36. Hyduke DR, Jarboe LR, Tran LM, Chou KJ, Liao JC: **Integrated network analysis identifies nitric oxide response networks and dihydroxyacid dehydratase as a crucial target in Escherichia coli.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104(20)**:8484-8489.
37. Jarboe LR, Hyduke DR, Tran LM, Chou KJ, Liao JC: **Determination of the Escherichia coli S-nitrosoglutathione response network using integrated biochemical and systems analysiS.** *The Journal of biological chemistry* 2008, **283(8)**:5148-5157.
38. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, Roth JR, Bäumler AJ: **Gut inflammation provides a respiratory electron acceptor for Salmonella.** *Nature* 2010, **467(7314)**:426-429.
39. Suthers PF, Zomorrodi A, Maranas CD: **Genome-scale gene/reaction essentiality and synthetic lethality analysis.** *Molecular systems biology* 2009, **5**:301.
40. Zimmermann GR, Lehar J, Keith CT: **Multi-target therapeutics: when the whole is greater than the sum of the parts.** *Drug discovery today* 2007, **12(1-2)**:34-42.
41. Barthelmes J, Ebeling C, Chang A, Schomburg I, Schomburg D: **BRENDA, AMENDA and FRENDA: the enzyme information system in 2007.** *Nucleic Acids Res* 2007, , **35 Database:** D511-514.
42. Garcia-del Portillo F, Nunez-Hernandez C, Eisman B, Ramos-Vivas J: **Growth control in the Salmonella-containing vacuole.** *Current opinion in microbiology* 2008, **11(1)**:46-52.
43. Shi L, Adkins JN, Coleman JR, Schepmoes AA, Dohnkova A, Mottaz HM, Norbeck AD, Purvine SO, Manes NP, Smallwood HS, Wang H, Forbes J, Gros P, Uzzau S, Rodland KD, Heffron F, Smith RD, Squier TC: **Proteomic analysis of Salmonella enterica serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages.** *J Biol Chem* 2006, **281(39)**:29131-29140.
44. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: **KEGG for linking genomes to life and the environment.** *Nucleic Acids Res* 2008, , **36 Database:** D480-484.
45. Fleming RM, Thiele I, Nasheuer HP: **Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to Escherichia coli.** *Biophys Chem* 2009, **145(2-3)**:47-56.
46. Varma A, Palsson BO: **Metabolic Flux Balancing: Basic concepts, Scientific and Practical Use.** *Nat Biotechnol* 1994, **12**:994-998.
47. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NER, Kang J, Hydruke D, Palsson BO: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0.** *Nat Prot.*

doi:10.1186/1752-0509-5-8
Cite this article as: Thiele *et al.*: A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella* Typhimurium LT2. *BMC Systems Biology* 2011 **5**:8.

2.2 Nutrient starvation limits *Salmonella* virulence during systemic infection

(Manuscript submitted)

Steeb B, Müller B, Burton N, Tienz P, Schlüter K, Busse D, Rabsch W, Biehlmaier O, Schmidt A, Farhan H, Mazé A, Bumann D

Summary:

Intracellular pathogens are dependent on nutrient supply from the host for survival and replication. However, comprehensive quantitative data about pathogen *in vivo* nutrition are lacking. This poor understanding of relevant *in vivo* growth conditions has led to major antimicrobial drug development failures.

Here, we determined the *in vivo* nutritional landscape of *S.* Typhimurium in a typhoid fever model. For this, we infected mice with nutrient utilization and auxotrophic mutants, to determine nutrient availability in *Salmonella* containing vacuoles (SCVs). Combined with literature data, we were able to identify 45 available host metabolites, with glycerol, fatty acids, N-acetylglucosamine, lactate and arginine representing major nutrients of *Salmonella in vivo*. In the next step, we used the *in vivo* replication rates of the *Salmonella* mutants to obtain quantitative nutrient flux data. For this, we refined the *Salmonella* metabolism reconstruction STMv1.0 with experimental data. The resulting *Salmonella in vivo* model of infection (STMv1.1) contained all detected nutrients and predicted *Salmonella* phenotypes described in literature with an accuracy of 90%. Furthermore, this model was supported by *ex vivo* proteome data. Thus, we presented for the first time a comprehensive quantitative report of *in vivo* nutrient utilization of a pathogen.

Our data revealed the paradoxical situation that the SCV is a qualitatively nutrient rich environment, which supported growth of all tested auxotrophic mutants. On the other hand, the inability to use one of the major nutrients led to a reduced growth rate, which indicated a quantitatively nutrient poor environment that limited *Salmonella in vivo* growth. Furthermore, we were able to show that an additional nutrient limitation by restricting host supply reduced the *in vivo* replication rate of *S.* Typhimurium. This indicated that inhibition of nutrient supply might be a promising approach to combat *Salmonella* infections and possibly other phagosomal pathogens such as *Legionella pneumophila* and *Mycobacterium tuberculosis*.

Statement of my work:

I generated a computational model of *Salmonella in vivo* metabolism (STMv1.1). For this, I fitted the metabolic reconstruction STMv1.0 to experimental data of nutrient utilization and auxotrophic mutant *in vivo* phenotypes. I determined quantitative *in vivo* flux estimates by using this infection model, which revealed *Salmonella* nutrient limitation during infection. I also participated in mutant construction (SL1344 *yabJ thiI*, SL1344 *pncB nadC*, SL1344 *pdxA STM0163*) and in *in vivo* virulence determination. I participated in writing the manuscript.

Nutrient starvation limits *Salmonella* virulence during systemic infection

Benjamin Steeb[1], Beatrice Müller[1], Neil Burton[1], Petra Tienz[1], Kai Schlüter[4], Dagmar Busse[5], Wolfgang Rabsch[5], Oliver Biehlmaier[2], Alexander Schmidt[3], Hesso Farhan[1], Alain Mazé[1], Dirk Bumann[1,4*]

[1]Focal Area Infection Biology, [2]Imaging Core Facility, and [3]Proteomics Core Facility, Biozentrum, University of Basel, CH-4056 Basel, Switzerland; [4]Junior Group "Mucosal Infections", Hannover Medical School, D-30625 Hannover, Germany; [5]Robert-Koch-Institute, D-38855 Wernigerode, Germany

Corresponding author:

Dirk Bumann

Biozentrum
Klingelbergstr. 50/70
CH-4056 Basel
Switzerland

Phone: +41 61 267 2382

E-mail: dirk.bumann@unibas.ch

Abstract

Host metabolites that support pathogen growth in tissues play an essential role in infectious diseases, but remain poorly characterized. Here, we identified and quantified relevant host nutrients in a *Salmonella* typhoid fever model, and used these data to build a quantitative genome-scale computational model of *Salmonella* nutrition, metabolism, and growth during infection. The data revealed availability of more than 45 diverse host metabolites that made *Salmonella* resilient against metabolic perturbation. However, *Salmonella* obtained these various metabolites in only scarce amounts that together just supported slow nutrient-limited growth. Nutrients were delivered to intracellular *Salmonella* through host cell endocytosis, and blocking this supply route further diminished *Salmonella* growth. In conclusion, these findings indicated a key role of qualitatively rich, yet quantitatively poor nutrient supply for *Salmonella* virulence, robustness, and control.

Infectious diseases are a major worldwide threat to human health (1). The situation is worsening because of rapidly rising antimicrobial resistance and insufficient development of new antibiotics. Most infectious diseases start with a few pathogenic organisms that invade host tissues, but disease symptoms develop only later when pathogens exploit host nutrients to grow to high tissue loads. Despite this crucial role of pathogen growth and nutrient supply, only few host nutrients that are relevant for individual pathogens have been identified (2-13), and comprehensive quantitative in vivo data are lacking for any pathogen. The poor understanding of relevant in vivo growth conditions can cause major antimicrobial drug development failures (14-17).

In this study, we investigated *Salmonella* nutrition and growth in a systemic mouse infection model mimicking human typhoid fever (18). This model is particularly suitable because of facile *Salmonella* genetics, availability of genome-scale metabolic reconstructions (19-21), extensive literature, and close similarities between *Salmonella* and the prime model organism *E. coli*.

Results

Salmonella access diverse nutrients in infected host tissues

To identify host nutrients that support *Salmonella* growth in infected mouse tissues during typhoid fever, we used *Salmonella* mutants with defects in utilization of specific nutrients. Most of these mutants retained normal in vivo growth rates indicating that 18 nutrients had limited relevance for *Salmonella* during systemic disease (Table S1,S2). On the other hand, mutants unable to utilize glycerol, fatty acids, N-acetylglucosamine, glucose, lactate, or arginine had significant growth defects suggesting that these six host nutrients can contribute to *Salmonella* in vivo growth (Fig. 1; see Table S2 for detailed analysis). Glucose was the only previously identified relevant nutrient in this disease model (11). The importance of these six nutrients was supported by high abundance of corresponding utilization enzymes in *Salmonella* purified from infected mice (Table S3, see below).

Similar mutant phenotypes were obtained for infected liver (Table S1) indicating that the same major nutrients supported *Salmonella* growth in two different host organs. In addition, small-scale experiments with genetically resistant 129/Sv mice carrying a functional *Slc11a1* allele (22) confirmed the importance of glycerol (or glycerol-3-phosphate) and N-acetyl-glucosamine for *Salmonella* growth (Fig. S1) suggesting similar *Salmonella* nutrition in susceptible (typhoid fever model) and resistant mice.

Nutrient utilization mutants revealed major nutrients that together explained some 85% of the *Salmonella* in vivo growth. However, this approach was unsuitable for minor nutrients because of limited sensitivity (some 3% detection threshold for in vivo growth defects). To circumvent this

limitation, we used auxotrophic *Salmonella* mutants with defects in biosynthesis of essential biomass components. These mutants could not grow unless they obtained the respective missing biomass component through external supplementation. Surprisingly, all tested auxotrophic mutants proliferated in vivo indicating that host tissues provided 16 additional metabolites including amino acids and several cofactors/cofactor precursors (Fig. 1; Tables S1,S2). As an example, *Salmonella pheA tyrA trpA* auxotrophic for phenylalanine, tyrosine, and tryptophan, retained full virulence indicating sufficient availability of all three aromatic amino acids to meet biomass requirements. This was surprising since *Salmonella aroA* which is unable to synthesize the aromatic amino acid precursor metabolite chorismate, is strongly attenuated (23). However, chorismate defects are not informative for aromatic amino acid availability since chorismate is also required for synthesis of ubiquinone, an essential biomass component that is unavailable in vivo (24).

We combined these data with previously reported phenotypes of *Salmonella* mutants and biomass requirements (Table S2) to obtain a comprehensive nutritional landscape for *Salmonella* in infected mouse spleen (Fig. 1). The data revealed *Salmonella* access to a remarkably wide range of diverse host nutrients during systemic infection. This qualitatively rich nutrition buffered numerous *Salmonella* metabolic defects (such as inactive amino acid biosynthesis), and thus contributed to the remarkable robustness of *Salmonella* metabolism during infection (24).

Qualitatively rich nutrition is typical for mammalian pathogens

To determine if these findings were representative for pathogen nutrition in general, we compared pathogen metabolic capabilities based on genome pathway annotations (25). We analyzed 153 different mammalian pathogen genomes for presence of 254 nutrient utilization pathways and 118 biosynthetic pathways (Fig. 2). Most pathogens shared the capability to utilize glycerol, fatty acids, various carbohydrates, nucleosides, and amino acids that could serve as N-sources (such as arginine), suggesting a general preference for the same nutrients that *Salmonella* used in the mouse typhoid fever model. Moreover, many pathogens lacked biosynthesis pathways for amino acids, nucleosides, and (pro)vitamins indicating that they - like *Salmonella* - could obtain diverse essential biomass components from their respective host environments. These data suggested that pathogen growth in infected mammalian tissues is supported by a large variety of commonly available host nutrients. The actual relevance of each nutrient might, however, vary for individual pathogens. As an example, *Mycobacterium tuberculosis* accesses fatty acids and a subset of host amino acids in infected mouse lung, while glycerol is not a major nutrient (17, 26-29).

A quantitative genome-scale model of *Salmonella* nutrition and growth

Our initial qualitative analysis revealed numerous host nutrients that supported *Salmonella* growth. As a next step towards a comprehensive quantitative understanding of *Salmonella* in vivo nutrition, we estimated uptake rates for these nutrients based on *Salmonella* mutant phenotypes and biomass requirements.

For auxotrophic mutants, the calculations were straightforward (Table S2). As an example, the proline auxotroph *Salmonella proC* had an in vivo generation time of 7 ± 2 h (compared to 6 ± 1.2 h for wildtype *Salmonella* (24)). This indicated that within 7 h, the host supplied enough proline to build a new *Salmonella* cell containing some 43.3 million proline molecules (19), which is equivalent to 1'700 ± 500 proline molecules s^{-1} per *Salmonella* cell. Analogous calculations yielded uptake rate estimates for 30 organic nutrients (Fig. 1; see Table S2 for detailed explanation). *Salmonella* wildtype growth rate, biomass requirements, and additional information yielded uptake rates for 15 inorganic nutrients.

Estimating uptake rates for the six major nutrients was more challenging because of their parallel metabolization to many different biomass components. To analyze this complex nutrient utilization, we used Flux-Balance Analysis (30) of a genome-scale computational model. Specifically, we updated a consensus reconstruction of the *Salmonella* metabolic network incorporating all available literature (21) (model STMv1.1 with 1277 *Salmonella* enzymes, 1822 metabolites, 2572 reactions; Table S4). We combined this reconstruction with 45 uptake rates for minor nutrients deduced from auxotrophic mutants (see above). We then determined uptake rates for major nutrients that were consistent with in vivo growth rates of informative *Salmonella* mutants. As an example, the growth defects of *Salmonella glpFK gldA glpT ugpB* (55 ± 20%) and related mutants (Tables S1,S2) revealed that glycerol contributed 45 ± 20 % of *Salmonella* in vivo growth at a generation time of 6 h. This required glycerol uptake at a rate of 33'000 ± 17'000 molecules s^{-1} per *Salmonella* cell. Analogous analysis provided 5 additional

uptake rates (see Table S2 for detailed explanation).

Combination of all determined nutrient uptake rates with the genome-scale *Salmonella* metabolism reconstruction yielded a comprehensive quantitative model of host nutrient supply, *Salmonella* nutrient conversion to biomass components, and overall *Salmonella* growth (Fig. 1; the model is available in SBML format at http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html).

Model validation with independent experimental data sets

The computational model was consistent with established extensive knowledge about *Salmonella* biochemistry, thermodynamics, and quantitative mutant phenotypes. To further validate the model, we compared large-scale predictions with independent experimental data sets.

The model predicted hundreds of metabolic fluxes that provided all required biomass components for *Salmonella* growth. Some reactions had highly variable flux predictions (“flux variability”; Fig. 3a) indicating that alternate metabolic states were compatible with the experimental constraints as previously observed in other systems (31). We determined the state with minimal total flux in the entire metabolic network (32, 33) which would minimize enzyme biosynthesis costs (34, 35) and enzyme crowding in the cytosol (36) (Figs. 1, 3a; an interactive map with detailed descriptions of all reactions is available at http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html).

To sustain these predicted metabolic fluxes, *Salmonella* would require appropriate amounts of the corresponding enzymes. To experimentally test these predicted requirements, we determined enzyme abundance in *Salmonella* purified from infected mouse spleen (see Material and Methods), and retrieved turnover numbers k_{cat} for *Salmonella* enzymes (or *E. coli* orthologues) from databases (37, 38) and additional literature. Based on these data, we calculated 242 maximal reaction rates $v_{max} = N \cdot k_{cat}$ (with N, number of enzyme molecules; k_{cat}, turnover number). (Fig. 3a; Table S3; an interactive map is available at http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html). Comparison with model predictions showed that each analyzed reaction had flux predictions within the experimentally determined feasible range (Fig. 3a).

For the specific “minimal total flux” state (see above), three reactions had unfeasibly high flux predictions. To avoid such discrepancies, we restricted all reaction rates with experimental data to ranges compatible with enzyme abundance and turnover numbers. The improved computational model yielded a fully consistent “minimal total flux” state with no change in overall flux distribution or growth rate. This consistency with large-scale experimental enzyme data supported the computational model.

Systematic analysis of enzyme essentiality for *Salmonella* in vivo growth provided additional experimentally testable predictions. Inactivation of most enzymes had no impact on predicted growth rate, and the few predicted essential enzymes were mostly involved in biosynthesis consistent with our experimental data (24) (an interactive phenotype prediction map is available

at
http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html). Detailed comparison with experimental semi-quantitative phenotypes of 799 enzyme defects revealed 90% prediction accuracy (Fig. 3b, large pie diagram; Supplementary Table 5 online) similar to accuracies achieved for the best computational models for *E. coli* in vitro cultures (39).

Some discrepancies between computational predictions and experimental data were probably due to experimental variation since 18 (9%) of the 193 genes with multiple experimental evidence, had conflicting data (Fig. 3b, small pie). This could reflect different experimental protocols; for example, low dose infections (some 1000 CFU) were used in most experiments (including this study), whereas high dose infections (some 10^6 CFU) were used in two large-scale studies (40, 41). In addition, some discrepancies likely reflect current model limitations including (i) neglect of enzyme functions unrelated to biomass generation (such as detoxification of reactive oxygen species), (ii) regulated isozyme expression, and (iii) differential enzyme/transporter substrate affinities. Taken together, largely consistent mutant phenotype predictions supported the computational model.

We also generated genome-scale models for *Salmonella enterica* serovars that cause human typhoid/paratyphoid fever (Typhi, Paratyphi A, Paratyphi C, and non-typhoidal *Salmonella* (NTS); model STY2 for *S*. Typhi strain Ty2 is available at
http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html). Most of these human-specific serovars have smaller

metabolic networks compared to serovar Typhimurium (which was the focus of this study) because of numerous gene deletions or inactivating mutations (42). Nevertheless, all models successfully produced biomass at comparable rates when supplied with nutrients available in infected mouse tissues. These results suggested that the mouse typhoid fever model represented a suitable approximation to *Salmonella* nutrition in human tissues. This was also supported by experimental data on tryptophan availability (43), *Salmonella* mutant phenotypes in human volunteers (44), in vitro nutrient utilization capabilities of serovar Typhi and Paratyphi A clinical isolates (see Supplementary Fig. 2 online), and generally similar *Salmonella*-host interactions in mouse typhoid fever and human systemic disease (18)

Taken together, experimental enzyme quantities and mutant/serovar phenotypes supported the computational model as an accurate quantitative genome-scale description of *Salmonella* nutrition, metabolism, and growth in infected tissues.

Nutrient starvation limits *Salmonella* growth

Infected mouse spleen provided a wide range of diverse host metabolites to *Salmonella*. Despite this qualitatively rich nutrition, individual uptake rates were rather low suggesting a quantitatively poor nutrition (Table S2). Even major nutrients such as glycerol or N-acetylglucosamine had supply rates that were some hundredfold lower compared to what would be required for fast *Salmonella* growth (Fig. 4a). Moreover, defects in utilization of just one out of six major nutrients affected *Salmonella* growth (Tables S1,S2) indicating the lack of any compensating surplus nutrients that could provide alternative

carbon/energy sources (Fig. 4b). Consistent with this qualitative argument, the computational model required overall nutrient limitation for consistent results. Specifically, when we forced the model to accommodate higher nutrient supply rates that would exceed *Salmonella* biomass requirements (i.e., overall nutrient excess), it became rapidly incompatible with experimental mutant data (Fig. 4c). All these data suggested quantitatively poor *Salmonella* nutrition that resulted in slow nutrient-limited in vivo growth.

To obtain independent experimental evidence, we used a cell culture infection model where *Salmonella* replicated intracellularly in macrophage-like cells mimicking conditions during systemic salmonellosis (45). In this cell culture model, extracellular metabolites can reach intracellular *Salmonella* and contribute to their nutrition (46-48). To test the impact of nutrient availability, we therefore added external glucose or mannitol to *Salmonella*-infected macrophage-like RAW 264.7 cells at 4 h post infection. Interestingly, both extracellular nutrients accelerated subsequent intracellular *Salmonella* growth. This was dependent on *Salmonella* glucose/mannitol-specific utilization capabilities (Fig. 4d) indicating that external glucose and mannitol directly contributed to *Salmonella* growth without prior host metabolization or other host effects. This was consistent with the fact that mannitol can not be metabolized by mammalian cells (49). Taken together, increasing nutrient availability accelerated *Salmonella* growth which confirmed the nutrient limitation of intracellular *Salmonella*.

Extracellular nutrient delivery through host cell endocytosis

Overall *Salmonella* nutrient limitation implied that mechanisms of host nutrient supply could be of central importance for *Salmonella* growth during systemic disease. In this disease, *Salmonella* reside intracellularly in so-called *Salmonella*-containing vacuoles (SCV's) (45). External nutrients could reach *Salmonella* in SCV's through (i) transport from the host cell cytosol across the SCV membrane, or (ii) host cell endocytosis followed by endosome-SCV fusion (48) (Fig. 5a).

The slow nutrient-limited *Salmonella* growth suggested limited access to nutrient-rich host cell cytosol (50). We tested this hypothesis using a *Salmonella* strain carrying an *uhpTp-gfp* transcriptional reporter fusion responsive to glucose-6-phosphate, a key metabolite of mammalian cytosol (50). This biosensor detected little glucose-6-phosphate in the microenvironment of wildtype *Salmonella* in infected spleen, whereas the same fusion readily responded in avirulent mutant *Salmonella sifA* that reside directly in the cytosol (51) (Fig. 5b). Poor access to glucose-6-phosphate for wildtype *Salmonella* was also consistent with wildtype growth of *Salmonella uhpT* defective for glucose-6-phosphate uptake (Table S1). Moreover, in cell culture infections external glucose reached *Salmonella* without prior phosphorylation (see above) although glucose entering mammalian cytosol is rapidly phosphorylated. Taken together, these data suggested an extracellular nutrient delivery route to *Salmonella* that bypassed the host cytosol.

One such delivery route could be host cell endocytosis followed by endosome-SCV fusion (Fig. 5a) as demonstrated in *Salmonella*-infected cell cultures (48). To detect delivery of endocytotic cargo to *Salmonella* in vivo, we

used a similar approach employing fluorescent dextran as a fluid-phase marker for endocytosis. Confocal microscopy of spleen cryosections obtained 2 h after intravenous dextran injection into *Salmonella*-infected mice, revealed many dextran-containing vesicles in close proximity to *Salmonella*. In some cases, dextran even appeared to be in direct contact with *Salmonella* (Fig. 5c, Supplemental Movies 1- 3) which might indicate vesicle-SCV fusion as previously observed in vitro (48), but spatial resolution in vivo was insufficient for conclusive interpretation.

To test the relevance of this endocytotic nutrient supply route for *Salmonella* growth, we disrupted dynamin-dependent endocytosis with the specific inhibitor dynasore (52), or macropinocytosis with the Na^+/H^+ exchanger (NHE) inhibitor EIPA (5-(N-ethyl-N-isopropyl)-amiloride) (53). In cell culture infections, we added inhibitors after *Salmonella* had already established their intracellular niche (4h post infection) to prevent interference with *Salmonella* host cell entry and endosome maturation. Under these conditions, dynasore had only weak effects on subsequent *Salmonella* growth, while EIPA largely prevented *Salmonella* growth (Fig. 5d) indicating a previously unknown crucial role of host cell macropinocytosis for supporting intracellular *Salmonella* growth. In vivo administration of dynasore had again no detectable effect on *Salmonella* growth, but oral EIPA administration to infected mice partially suppressed *Salmonella* growth in infected spleen (Fig. 5e).

Taken together, these data were compatible with a model in which intracellular *Salmonella* nutrition depends on host cell delivery of extracellular nutrients through endocytosis. Macropinocytosis seemed to be particularly

relevant in the in vitro model. In vivo additional EIPA-insensitive delivery routes such as interaction with secretory pathways (54) might be involved. On the other hand, EIPA is a pleiotropic inhibitor of Na^+/H^+ exchangers involved in many biological processes including inflammation (55) which could also limit its in vivo efficacy for infection control. Further studies are required to clarify this issue.

Discussion

Host nutrients are essential for pathogen in vivo proliferation and disease progression, but comprehensive quantitative data are lacking. This study used complementary experimental and computational approaches to determine the relationship between *Salmonella* nutrition and growth in a mouse typhoid fever model.

Salmonella mutant in vivo phenotypes revealed access to an unexpected large number of chemically diverse host nutrients including lipids, carbohydrates, amino acids, nucleosides, and various (pro)vitamins. This diverse nutrition made *Salmonella* resilient against numerous metabolic perturbations. Genome comparisons revealed that *Salmonella* shared this qualitatively rich nutrition with many other mammalian pathogens. This might imply robust pathogen metabolism as a general challenge for antimicrobial therapy of infectious diseases.

The common availability of specific host metabolites likely reflects the general biochemical composition of mammalian tissues. Interestingly, some of the prototypical nutrients are predominantly present as part of high molecular weight compounds such as glycans/glycoproteins (GlcNAc), proteins (most amino acids), or lipids (glycerol, fatty acids) suggesting that macromolecule hydrolysis might be an important aspect of pathogen nutrition in infected tissues. Indeed, many pathogens express hydrolases as part of their virulence program, but host hydrolases might also be involved. Further research is needed to clarify this issue.

In addition to this qualitative analysis, we also obtained quantitative in vivo estimates for more than 45 individual nutrient uptake rates. Combination

of these rates with a genome-scale reconstruction of the *Salmonella* metabolic network yielded a quantitative computational model of *Salmonella* nutrient uptake, metabolization, and growth in infected host tissues. This model describes how *Salmonella* uses its metabolic network involving hundreds of enzymes to convert diverse host nutrients to biomass components for growth, and explains virulence phenotypes of more than 700 metabolic defects. Consistency of these predictions with independent large-scale experimental data suggested that the model accurately captured most relevant aspects of *Salmonella* nutrition and growth during infection.

Integrated quantitative analysis revealed that the many diverse nutrients were available in only scarce amounts that together just supported slow nutrient-limited *Salmonella* growth. *Salmonella* thus seemed to face a paradoxical situation with qualitatively rich, but quantitatively poor nutrition. Nutrient starvation limited *Salmonella* in vivo growth and thus represented a partially effective host mechanism to control *Salmonella*. Such an important impact of nutrient starvation was initially surprising since *Salmonella* infection elicited a potent inflammatory host response (see Supplementary Fig. 3 online) that could restrict *Salmonella* growth through various antibacterial mechanisms unrelated to metabolism. However, *Salmonella* has evolved multiple defense systems that effectively subvert and annihilate such host antibacterial effector mechanisms (56, 57) while poor nutrient supply might be more difficult to circumvent.

Partial *Salmonella* control through nutrient starvation suggested that residual host nutrient supply could be crucial for *Salmonella* virulence and disease progression. Only one membrane separates intracellular *Salmonella*

from nutrient-rich host cell cytosol, but this membrane was apparently rather effective to limit *Salmonella* access to abundant cytosolic metabolites. Instead, *Salmonella* obtained external nutrients through host cell endocytosis followed by fusion of endocytotic vesicles with the *Salmonella*-containing vacuole. Blocking endocytosis diminished *Salmonella* growth in in vitro cell culture infection models suggesting that manipulation of nutrient supply could offer a potential strategy for infection control. Indeed, similar pharmacological intervention partially inhibited *Salmonella* growth in the murine disease model, but specific, safe endocytosis inhibitors with increased in vivo efficacy have yet to be developed to rigorously test this strategy. Targeting host nutrient supply might also be considered for other slowly growing intracellular pathogens such as *Legionella pneumophila* and *Mycobacterium tuberculosis* that exploit host nutrient delivery pathways (58, 59).

In summary, this study provided a comprehensive quantitative in vivo analysis of nutrient supply and conversion into new biomass for an important pathogen. The findings revealed a key role of host-pathogen metabolic interactions for pathogen growth, disease progression, and development of novel infection control strategies.

Materials and Methods

Full methods and associated references are described in SI Materials and Methods. *Salmonella* mutants were generated by lambda red-recombinase mediated allelic replacement (60) followed by general transduction using phage P22 int (61). Mice were infected intravenously with 500-2000 CFU *Salmonella. Salmonella* were purified from infected mouse spleen using flow cytometry as described (24). Enzyme quantities were determined by shot-gun proteomics using heavy isotope-labeled peptides as internal references (62). The computational model of *Salmonella* nutrition and growth was based on an updated genome-scale reconstruction (21). Fluxes and mutant phenotypes were predicted using Flux-Balance Analysis (FBA) with the COBRA toolbox (33).

Acknowledgements

We thank Petra Spröte and Kathrin Schemmer for technical help with the dextran in vivo experiment, Nicole Freed for providing an *uhpTp-gfp* fusion, and Tomer Altman and Peter Karp for providing pathway annotations for bacterial genomes. Di.B. and co-workers are funded by the Swiss National Science Foundation (31003A-121834), SystemsX (RTD project BattleX) and Deutsche Forschungsgemeinschaft (SFB641-A9, SPP1316 Bu971/6).

References

1. Lopez AD, Mathers CD, Ezzati M, Jamison DT, & Murray CJ (2006) Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 367:1747-1757.
2. Chang DE*, et al.* (2004) Carbon nutrition of Escherichia coli in the mouse intestine. *Proc.Natl.Acad.Sci.U.S.A* 101:7427-7432.
3. Liu J, Istvan ES, Gluzman IY, Gross J, & Goldberg DE (2006) Plasmodium falciparum ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proc Natl Acad Sci U S A* 103:8840-8845.
4. Munoz-Elias EJ & McKinney JD (2006) Carbon metabolism of intracellular bacteria. *Cell Microbiol* 8:10-22.
5. Hofreuter D, Novik V, & Galan JE (2008) Metabolic diversity in Campylobacter jejuni enhances specific tissue colonization. *Cell Host Microbe* 4:425-433.
6. Olszewski KL*, et al.* (2009) Host-parasite interactions revealed by Plasmodium falciparum metabolomics. *Cell Host Microbe* 5:191-199.
7. Alteri CJ, Smith SN, & Mobley HL (2009) Fitness of Escherichia coli during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLoS Pathog* 5:e1000448.
8. Eisenreich W, Dandekar T, Heesemann J, & Goebel W (2010) Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nat Rev Microbiol* 8:401-412.
9. Polonais V & Soldati-Favre D (2010) Versatility in the acquisition of energy and carbon sources by the Apicomplexa. *Biol Cell* 102:435-445.
10. Marrero J, Rhee KY, Schnappinger D, Pethe K, & Ehrt S (2010) Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for Mycobacterium tuberculosis to establish and maintain infection. *Proc Natl Acad Sci U S A* 107:9819-9824.
11. Bowden SD, Rowley G, Hinton JC, & Thompson A (2009) Glucose and glycolysis are required for the successful infection of macrophages and mice by Salmonella enterica serovar typhimurium. *Infect Immun* 77:3117-3126.
12. Winter SE*, et al.* (2010) Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* 467:426-429.
13. Thiennimitr P*, et al.* (2011) Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci U S A* 108:17480-17485.
14. Arias CA & Murray BE (2009) Antibiotic-resistant bugs in the 21st century--a clinical super-challenge. *N Engl J Med* 360:439-443.
15. Payne DJ, Gwynn MN, Holmes DJ, & Pompliano DL (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* 6:29-40.

16. Brinster S*, et al.* (2009) Type II fatty acid synthesis is not a suitable antibiotic target for Gram-positive pathogens. *Nature* 458:83-86.

17. Pethe K*, et al.* (2010) A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. *Nat Commun* 1:57.

18. Tsolis RM, Xavier MN, Santos RL, & Baumler AJ (2011) How to become a top model: The impact of animal experimentation on human Salmonella disease research. *Infect Immun* 79:1806-1814.

19. Raghunathan A, Reed J, Shin S, Palsson B, & Daefler S (2009) Constraint-based analysis of metabolic capacity of Salmonella typhimurium during host-pathogen interaction. *BMC Syst Biol* 3:38.

20. AbuOun M*, et al.* (2009) Genome scale reconstruction of a Salmonella metabolic model: comparison of similarity and differences with a commensal Escherichia coli strain. *J Biol Chem* 284:29480-29488.

21. Thiele I*, et al.* (2011) A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2. *BMC Syst Biol* 5:8.

22. Bellamy R (1999) The natural resistance-associated macrophage protein and susceptibility to intracellular pathogens. *Microbes.Infect.* 1:23-27.

23. Hoiseth SK & Stocker BA (1981) Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines. *Nature* 291:238-239.

24. Becker D*, et al.* (2006) Robust Salmonella metabolism limits possibilities for new antimicrobials. *Nature.* 440:303-307.

25. Caspi R*, et al.* (2009) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 38:D473-479.

26. Hondalus MK*, et al.* (2000) Attenuation of and protection induced by a leucine auxotroph of Mycobacterium tuberculosis. *Infect Immun* 68:2888-2898.

27. Smith DA, Parish T, Stoker NG, & Bancroft GJ (2001) Characterization of auxotrophic mutants of Mycobacterium tuberculosis and their potential as vaccine candidates. *Infect Immun* 69:1142-1150.

28. Pavelka MS, Jr., Chen B, Kelley CL, Collins FM, & Jacobs Jr WR, Jr. (2003) Vaccine efficacy of a lysine auxotroph of Mycobacterium tuberculosis. *Infect Immun* 71:4190-4192.

29. Senaratne RH*, et al.* (2007) Vaccine efficacy of an attenuated but persistent Mycobacterium tuberculosis cysH mutant. *J Med Microbiol* 56:454-458.

30. Oberhardt MA, Palsson BO, & Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 5:320.

31. Reed JL & Palsson BO (2004) Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated

reaction subsets that comprise network states. *Genome Res* 14:1797-1805.

32. Holzhutter HG (2004) The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *Eur J Biochem* 271:2905-2922.

33. Becker SA*, et al.* (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* 2:727-738.

34. Akashi H & Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. *Proc.Natl.Acad.Sci.U.S.A* 99:3695-3700.

35. Dekel E & Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436:588-592.

36. Beg QK*, et al.* (2007) Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proc Natl Acad Sci U S A* 104:12663-12668.

37. Scheer M*, et al.* (2011) BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* 39:D670-676.

38. Keseler IM*, et al.* (2011) EcoCyc: a comprehensive database of Escherichia coli biology. *Nucleic Acids Res* 39:D583-590.

39. Orth JD*, et al.* (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism-2011. *Mol Syst Biol* 7:535.

40. Santiviago CA*, et al.* (2009) Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLoS Pathog* 5:e1000477.

41. Chaudhuri RR*, et al.* (2009) Comprehensive identification of Salmonella enterica serovar typhimurium genes required for infection of BALB/c mice. *PLoS Pathog* 5:e1000529.

42. Holt KE*, et al.* (2009) Pseudogene accumulation in the evolutionary histories of Salmonella enterica serovars Paratyphi A and Typhi. *BMC Genomics* 10:36.

43. Virgilio R & Cordano AM (1981) Naturally occurring prototrophic strains of Salmonella typhi. *Can.J.Microbiol.* 27:1272-1275.

44. Bumann D, Hueck C, Aebischer T, & Meyer TF (2000) Recombinant live Salmonella spp. for human vaccination against heterologous pathogens. *FEMS Immunol.Med.Microbiol.* 27:357-364.

45. Haraga A, Ohlson MB, & Miller SI (2008) Salmonellae interplay with host cells. *Nat Rev Microbiol* 6:53-66.

46. Leung KY & Finlay BB (1991) Intracellular replication is essential for the virulence of Salmonella typhimurium. *Proc.Natl.Acad.Sci.U.S.A* 88:11470-11474.

47. Buchmeier NA & Libby SJ (1997) Dynamics of growth and death within a Salmonella typhimurium population during infection of macrophages. *Can.J Microbiol* 43:29-34.

48. Drecktrah D, Knodler LA, Howe D, & Steele-Mortimer O (2007) Salmonella trafficking is defined by continuous dynamic interactions with the endolysosomal system. *Traffic* 8:212-225.

49. Livesey G (2003) Health potential of polyols as sugar replacers, with emphasis on low glycaemic properties. *Nutr Res Rev* 16:163-191.

50. Wishart DS*, et al.* (2009) HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* 37:D603-610.

51. Beuzon CR*, et al.* (2000) Salmonella maintains the integrity of its intracellular vacuole through the action of SifA. *EMBO J.* 19:3235-3249.

52. Macia E*, et al.* (2006) Dynasore, a cell-permeable inhibitor of dynamin. *Dev Cell* 10:839-850.

53. Koivusalo M*, et al.* (2010) Amiloride inhibits macropinocytosis by lowering submembranous pH and preventing Rac1 and Cdc42 signaling. *J Cell Biol* 188:547-563.

54. Kuhle V, Abrahams GL, & Hensel M (2006) Intracellular Salmonella enterica redirect exocytic transport processes in a Salmonella pathogenicity island 2-dependent manner. *Traffic* 7:716-730.

55. Pedersen SF (2006) The Na+/H+ exchanger NHE1 in stress-induced signal transduction: implications for cell proliferation and cell death. *Pflugers Arch* 452:249-259.

56. Valdez Y, Ferreira RB, & Finlay BB (2009) Molecular mechanisms of Salmonella virulence and host resistance. *Curr Top Microbiol Immunol* 337:93-127.

57. Aussel L*, et al.* (2011) Salmonella detoxifying enzymes are sufficient to cope with the host oxidative burst. *Mol Microbiol* 2011:1365-2958.

58. Sturgill-Koszycki S & Swanson MS (2000) Legionella pneumophila replication vacuoles mature into acidic, endocytic organelles. *J Exp Med* 192:1261-1272.

59. Backus KM*, et al.* (2011) Uptake of unnatural trehalose analogs as a reporter for Mycobacterium tuberculosis. *Nat Chem Biol* 7:228-235.

60. Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc.Natl.Acad.Sci.U.S.A* 97:6640-6645.

61. Thierauf A, Perez G, & Maloy AS (2009) Generalized transduction. *Methods Mol Biol* 501:267-286.

62. Maier T*, et al.* (2011) Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol Syst Biol* 7:511.

Figure legends

Figure 1: A quantitative genome-scale model of *Salmonella* nutrition, metabolism, and growth in infected mouse spleen. This schematic map shows available host nutrients, their respective uptake rates represented by color and font size, and their metabolization to new biomass (see text and Supplementary Tables 1, 2, 3 online for detailed explanation, quantitative values, and statistical analysis). An interactive map with detailed annotation of all reactions is available at http://www.biozentrum.unibas.ch/personal/bumann/supplemental_information/steeb_et_al/index.html.

Figure 2: A general nutrition pattern for mammalian pathogens. **Left**) Frequency of 254 nutrient utilization pathways in genomes of 153 mammalian pathogens (excluding all *Salmonella* serovars). A frequency of 1 indicates pathway presence in all pathogens. Data were based on pathway annotations available in MetaCyc (25). Degradation pathways for nutrients that support *Salmonella* in mouse spleen (filled circles) were highly overrepresented among pathogen genomes ($P < 0.001$; Mann-Whitney U test) suggesting similar nutritional preferences. **Right**) Depletion frequency of 118 biosynthesis pathways in mammalian pathogens. The values represent differences in pathway frequency in sets of 153 pathogens and 316 environmental bacteria. As an example, a biosynthesis pathway present in 30% of mammalian pathogens and 50% of environmental bacteria would be assigned a “depletion frequency” of 0.3 - 0.5 = -0.2. Biosynthesis pathways for biomass components that *Salmonella* could obtain from the host (filled circles) were selectively

depleted among pathogen genomes (P< 0.001; Mann-Whitney U test) suggesting similar host supplementation patterns.

Figure 3: Large-scale experimental data validate the computational model (Figure 1). **a**) Feasibility of predicted reaction rates. For each reaction, the range of flux rates compatible with data for *Salmonella* nutritional mutants was predicted using Flux-Balance Analysis. The circles represent the flux state with minimal total flux. These data were compared to feasible reaction rate estimates calculated from experimental enzyme copy numbers and kinetic parameters (see Supplementary Table 3 online). The shaded area represents unfeasible fluxes. All reactions had predicted fluxes within the feasible ranges. **b**) Mutant phenotype predictions. Comparison of model predictions with 799 experimental *Salmonella* mutant phenotypes revealed 9.5% discrepancies (large circle, red area). This might in part reflect experimental inconsistencies that were revealed by conflicting results in 8.8% of 193 genes with multiple experimental studies (small circle, red area).

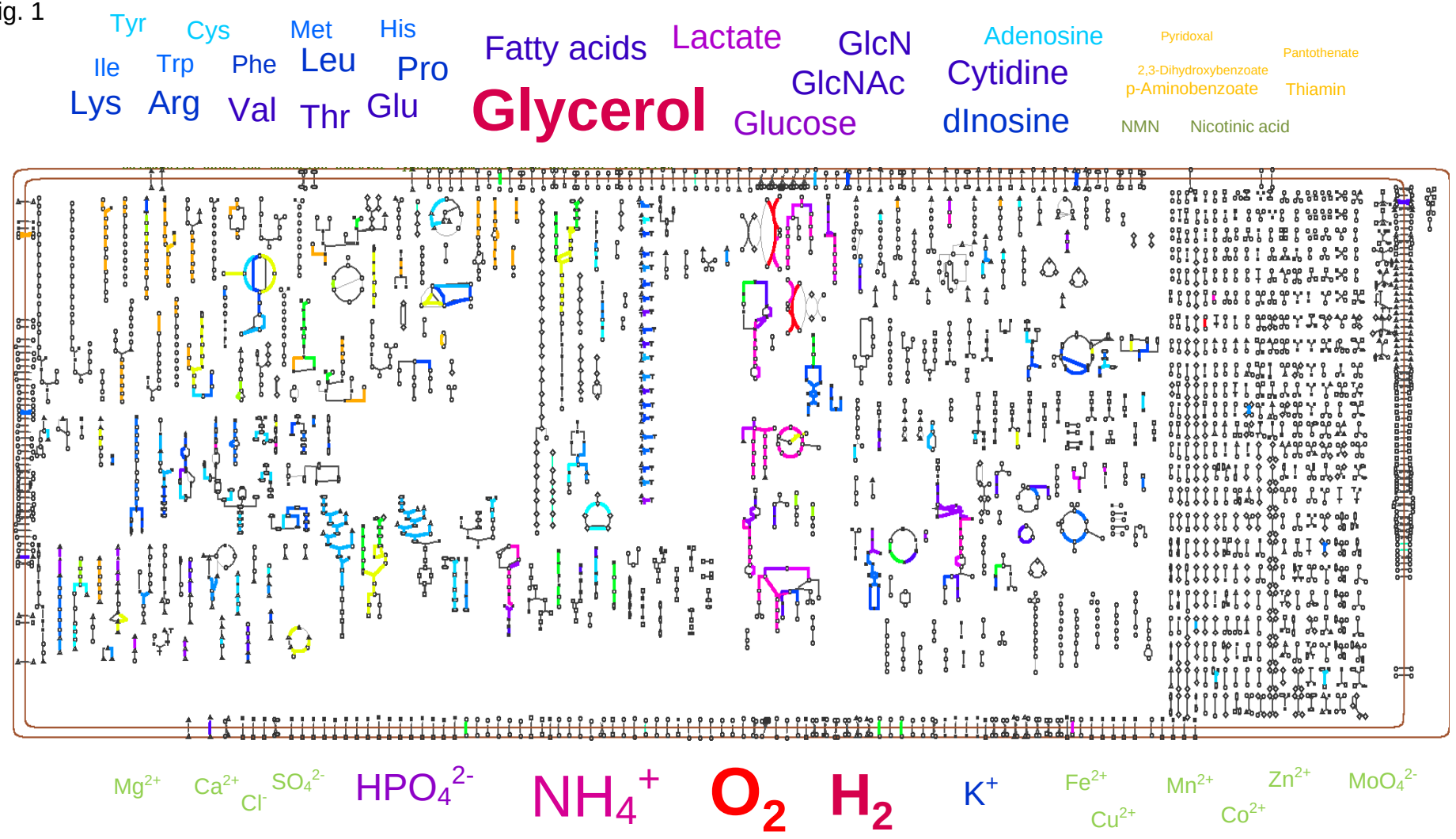
Figure 4: Nutrient starvation limits *Salmonella* in vivo growth. **a)** *Salmonella* uptake rates that are consistent with in vivo growth rates of informative *Salmonella* mutants for six major nutrients (open circles) are compared to uptake rates for the same nutrients that would support rapid *Salmonella* growth with a generation time of 30 min. (filled circles). **b**) Schematic comparison of slow *Salmonella* growth with limiting or excess amounts of nutrients. In case of nutrient-limitation, defects in utilization of any relevant

nutrient diminish growth. This scenario is consistent with experimental data. In case of nutrient excess, *Salmonella* growth is restricted by some nutrient-unrelated host inhibitory mechanism that prevents *Salmonella* from exploiting all available nutrients. A defect in utilization of one of the various nutrients would have little impact since *Salmonella* can readily switch to other excess nutrients. This scenario is incompatible with our experimental data. **c**) Flux-Balance Analysis of *Salmonella* mutant data for various nutrient excess scenarios. Mean square errors for predicted vs. experimental *Salmonella* mutant growth defects are shown. Scenarios with nutrient excess result in model predictions that are inconsistent with experimental data. **d**) Increasing external nutrient availability in an in vitro cell culture model accelerates intracellular *Salmonella* growth in macrophage-like cells, and this depends on specific *Salmonella* nutrient utilization capabilities (open symbols, 0.5 g l^{-1} glucose; filled grey symbols, 1 g l^{-1} glucose; filled black symbols, 0.5 g l^{-1} glucose 0.5 g l^{-1} mannitol; circles, wildtype *Salmonella*; upward triangles, *Salmonella ptsG manX galP mglB*, deficient for high-affinity glucose transport; downward triangles, *Salmonella mtlAD*, deficient for high-affinity mannitol transport and degradation). Colony-forming units (CFU) at 10 h post infection for triplicate wells containing 300’000 RAW 264.7 cells are shown.

Figure 5: Host cell endocytosis provides a nutrient supply route for intracellular *Salmonella*. **a**) Schematic representation of alternative host supply routes (1, transport from host cell cytosol across the vacuolar membrane; 2, transfer of endocytotic cargo by fusion of the *Salmonella*-containing vacuole with late endosomes). **b**) Poor access of intracellular

Salmonella to a key cytosolic metabolite. Wildtype *Salmonella* carrying an *uhpTp-gfp* fusion responsive to glucose-6-phosphate showed weak in vivo activity (blue) in contrast to a *Salmonella sifA* mutant (red) residing in host cell cytosol (51) (the grey curve represents GFP fluorescence of an in vitro culture without glucose-6-phosphate). Similar data were obtained in two independent experiments. **c**) Topology of fluid-phase marker dextran (green) and *Salmonella* (red) in infected mouse spleen. Confocal stacks are shown as maximum intensity projections (upper panels) and three-dimensional surface renderings (lower panels). The scale bars represent 1 µm. See also Movies 1-3 available online. **d**) Intracellular *Salmonella* growth in macrophage-like cells in vitro in control wells, or in presence of 30 µM dynasore or 30 µM EIPA. The macropinocytosis inhibitor EIPA blocks intracellular *Salmonella* growth. Data represent CFU values from triplicate wells. **e**) *Salmonella* load in spleen of control mice and mice treated with EIPA. The data represent CFU values of individual mice from two independent experiments.

Fig. 1



Uptake / reaction rates in molecules s^{-1} per *Salmonella*

Fig. 2

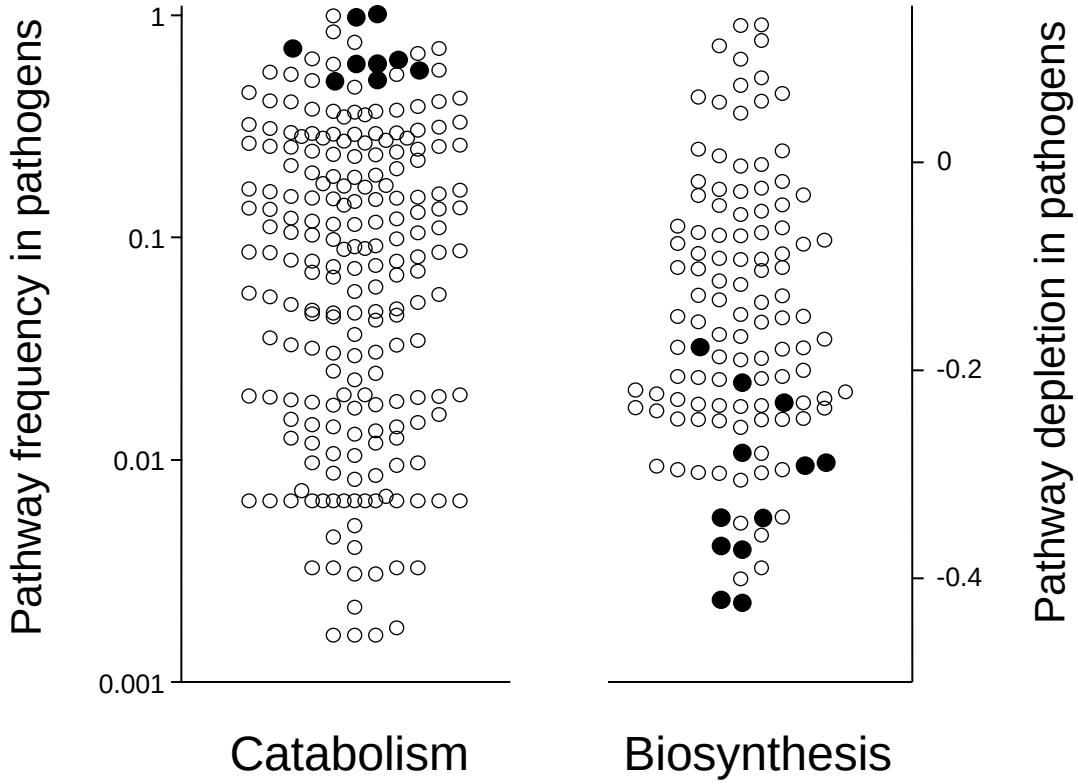

Pathway frequency in pathogens
1
0.1
0.01
0.001
Catabolism
Biosynthesis
0
-0.2
-0.4
Pathway depletion in pathogens

Fig. 3

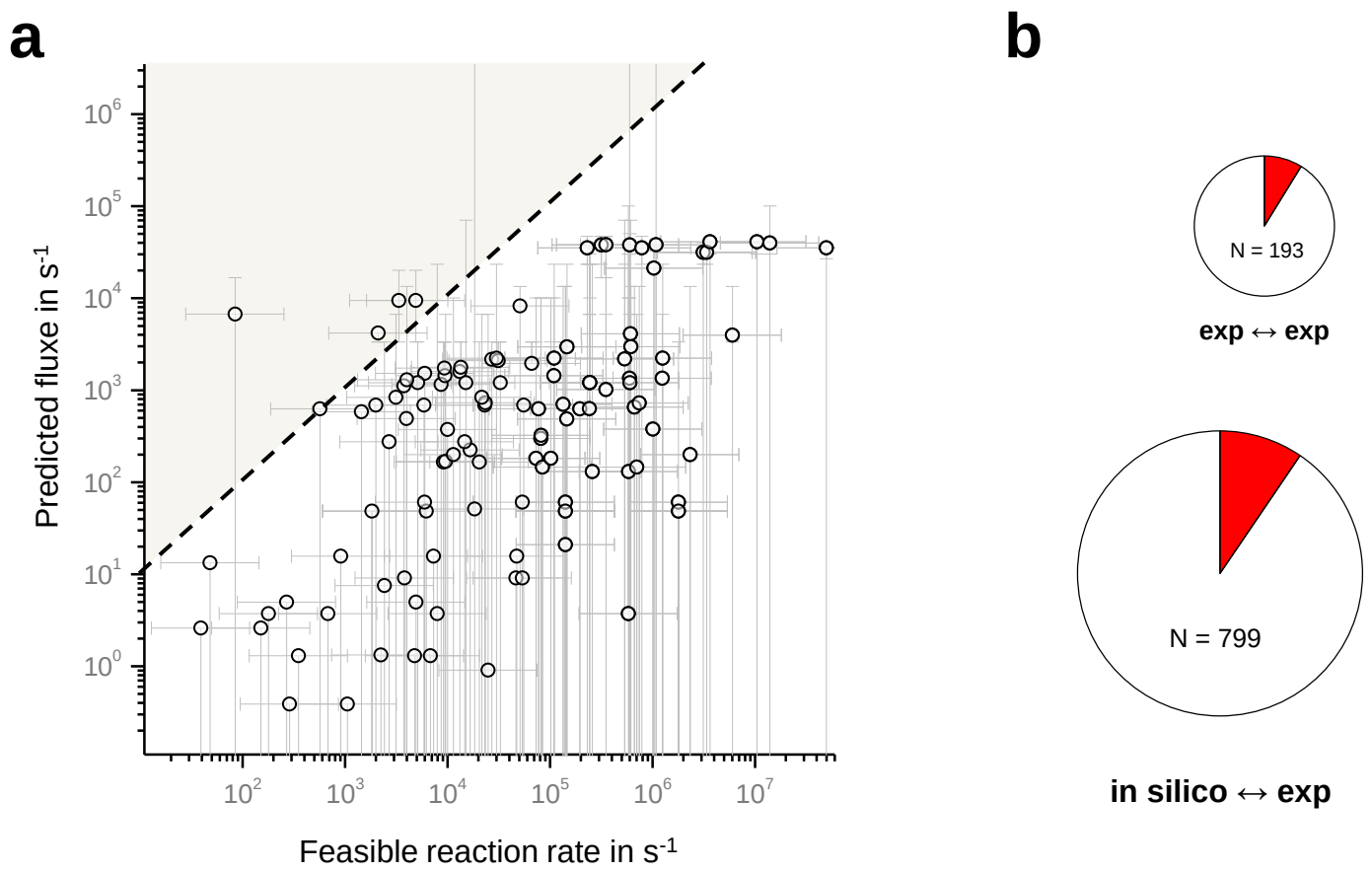

a
b
Predicted fluxe in s-1
Feasible reaction rate in s-1
N = 193
exp ↔ exp
N = 799
in silico ↔ exp

Fig. 4

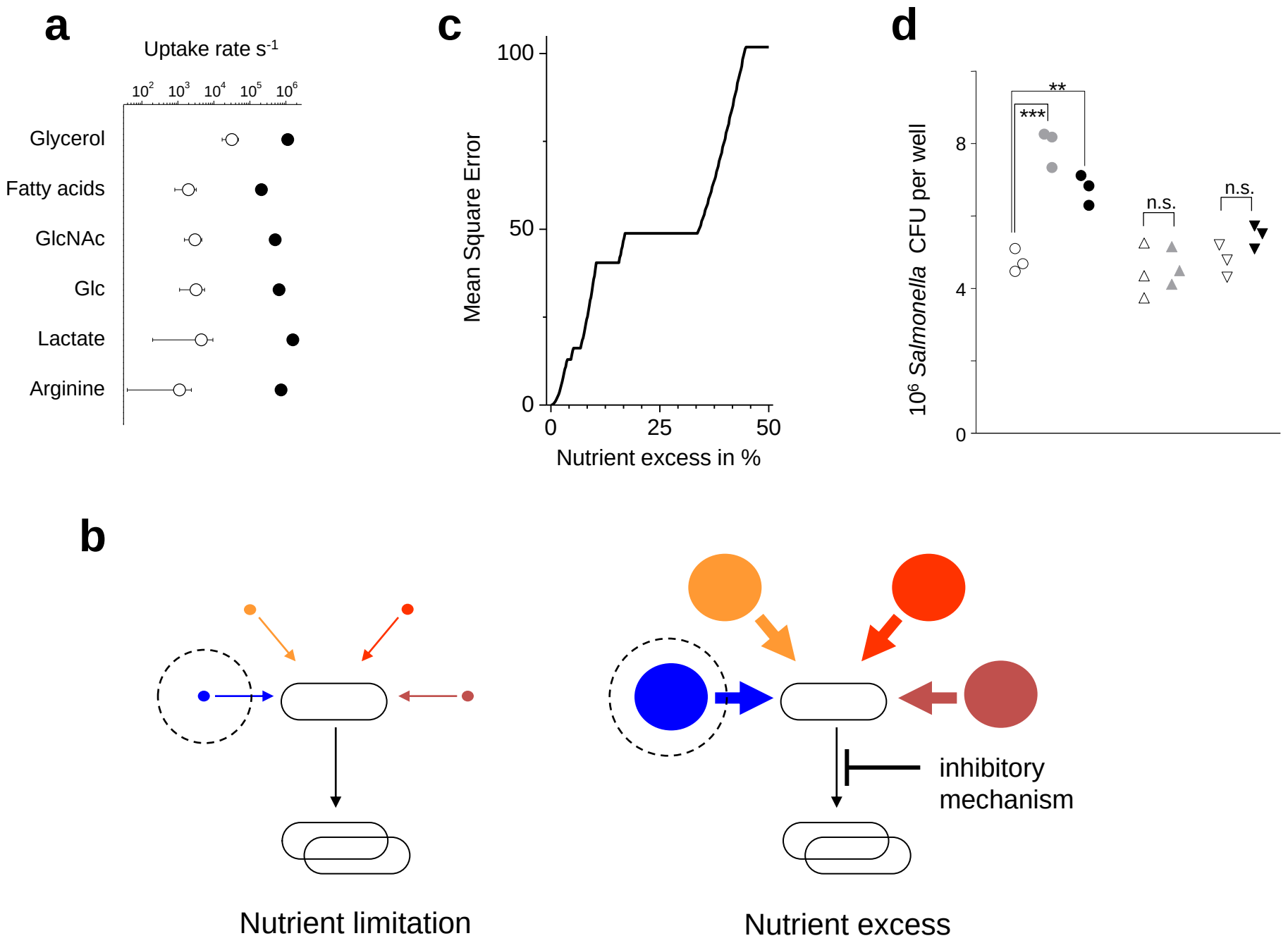
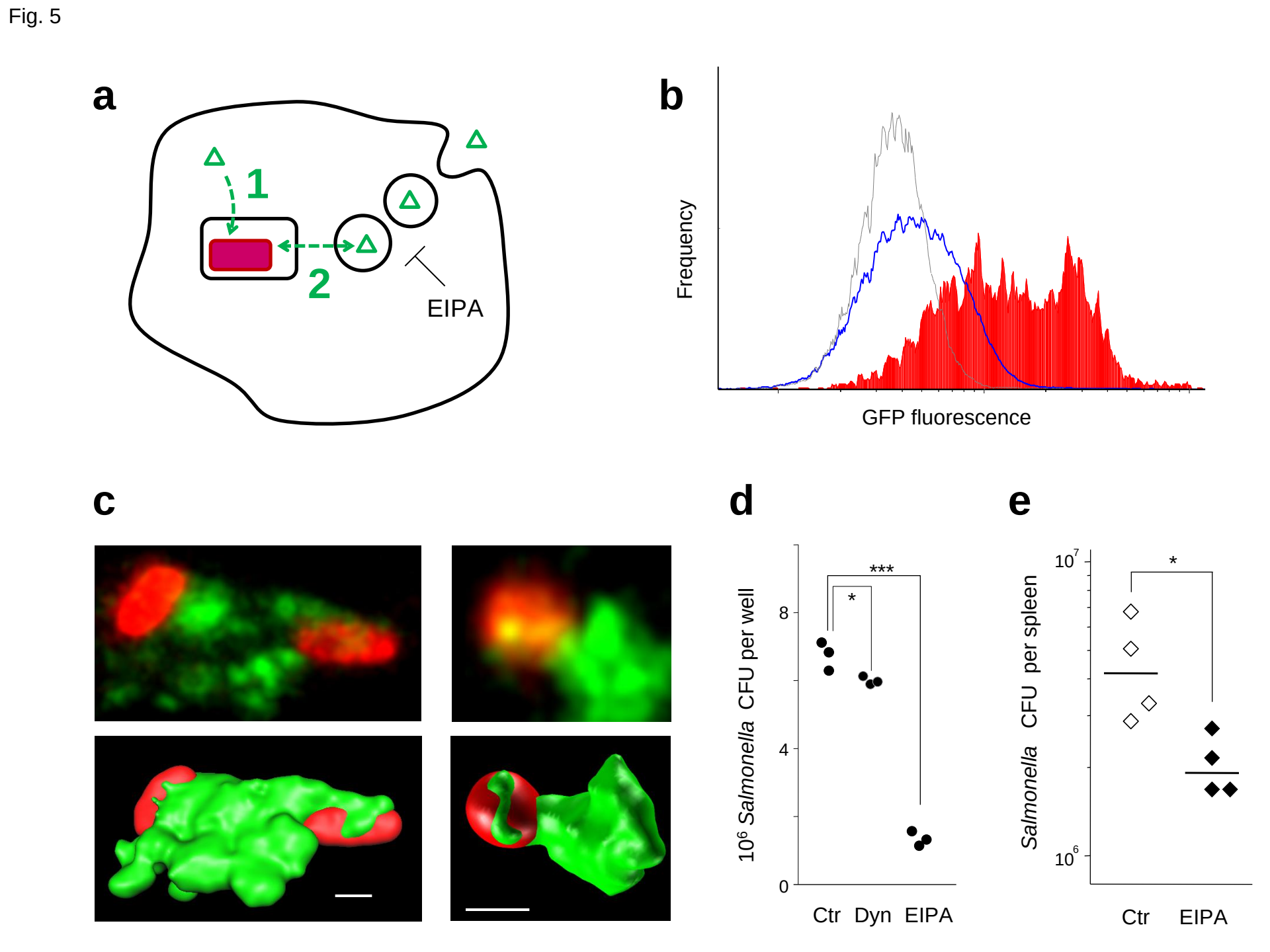

a
Uptake rate s-1
10^2 10^3 10^4 10^5 10^6
Glycerol
Fatty acids
GlcNAc
Glc
Lactate
Arginine
c
Mean Square Error
100
50
0
0 25 50
Nutrient excess in %
d
10^6 Salmonella CFU per well
**
***
n.s.
n.s.
8
4
0
b
inhibitory mechanism
Nutrient limitation
Nutrient excess

Fig. 5


a
1
2
EIPA
b
Frequency
GFP fluorescence
c
d
10^6 *Salmonella* CFU per well
8
4
0
***
*
Ctr
Dyn
EIPA
e
Salmonella CFU per spleen
10^7
10^6
*
Ctr
EIPA

SI Text

SI Materials and Methods. *Bacterial genetics.* *Salmonella* mutants were constructed by lambda red-recombinase mediated allelic replacement (1) followed by general transduction using phage P22 int (2). Resistance cassettes flanked with FRT sites were removed using FLP recombinase (1). To generate *Salmonella* nutrient utilization mutants, we preferentially disrupted nutrient uptake instead of deleting internal catabolic enzymes to minimize accumulation of potentially toxic intermediates such as phosphorylated carbohydrates (3, 4) that could complicate mutant phenotype interpretation. We constructed a series of *Salmonella* nutrient uptake mutants. Because *Salmonella* has multiple high-affinity uptake systems for many nutrients, we had to engineer strains containing multiple mutations. This made it difficult to complement mutants to exclude effects of secondary mutations. To validate major phenotypes we instead independently reconstructed two substantially attenuated mutants (*glpF glpK gldA glpT ugpB*; *nagE manX)*. In both cases, identical in vivo phenotypes were observed (data not shown). Phenotypes of auxotrophic mutants were validated using M9 minimal media containing appropriate supplements. A plasmid carrying a *uhpTp-gfp* fusion was transformed into SL1344 *sifB*::*cherry* to detect all *Salmonella* based on red fluorescence, and *Salmonella* subpopulations with active reporter constructs based on green fluorescence. Strains were cultivated on Lennox LB medium containing 90 μg ml^{-1} streptomycin, 50 μg ml^{-1} kanamycin, 20 μg ml^{-1} chloramphenicol, and/or 100 μg ml^{-1} ampicillin. Media were supplemented for auxotrophs as needed.

Mouse infections. We infected female, 8 to 12 weeks old BALB/c mice intravenously with 500-2000 CFU *Salmonella* from late exponential LB cultures. For some experiments, we also used female, 8-12 weeks old 129/Sv mice. In some cases, we administered EIPA in the drinking water (20 µg ml^{-1}) throughout the infection period. Four days (five days for 129/Sv) post-infection, mice were sacrificed and bacterial loads in spleen and liver were determined by plating of tissue homogenates treated with 0.3% Triton Tx-100. In competitive infections, wildtype and mutant *Salmonella* carrying different antibiotic resistance markers were mixed before administration. Individual strain tissue loads were determined by replica plating on selective media and competitive indices (CI=output ratio/input ratio) were calculated. For BALB/c mice we converted CI values to growth rates based on a 6 h in vivo generation time (5). For 129/Sv mice, we present the data in Fig. S1 as competitive indices since information on in vivo generation time was lacking. Statistical significance was analyzed using t-test (a parametric test was appropriate based on the normal distribution of growth rates (or log-transformed CI values) (5). Our experiments involved a large set of strains. To avoid the multiple comparison problem, we used the Benjamini-Hochberg false discovery rate (FDR) approach (6).

Flow Cytometry. Samples were analyzed using a Fortezza flow cytometer equipped with 488 nm and 561 nm lasers (BD Biosciences). For *Salmonella* ex vivo purification, *Salmonella sifB::gfp* (7) were sorted from infected mouse spleen as described (5) using a FacsAria III sorter equipped with 488 nm and 561 nm lasers (BD Biosciences). Proteome changes were minimized by

preventing de novo synthesis with 170 μM chloramphenicol and delaying proteolysis by maintaining the samples at 0-4°C. Our previous results suggested that these conditions were largely effective to preserve the in vivo *Salmonella* proteome during purification (5).

Enzyme quantification using mass spectrometry-based proteomics. Ex vivo sorted *Salmonella* were lysed and treated with trypsin. After adding a mix of heavy labeled reference AQUA peptides (Table S6), the peptides were purified using solid phase extraction (Macrospin columns, Harvard Apparatus) and analyzed by directed LC-MS/MS using an LTQ-Orbitrap-Velos instrument (Thermo-Fischer Scientific). In order to increase the number of protein identifications, MS-sequencing was focused on previously identified peptides from *Salmonella* using a recently developed inclusion list driven workflow (8). Peptides and proteins were searched against the SL1344 genome sequence (ftp://ftp.sanger.ac.uk/pub/pathogens/Salmonella/) including 206 frequently observed contaminants, all mouse entries from SwissProt (Version 57.12) and all sequences in reversed order (total 42502 entries) using the Mascot search algorithm. The false discovery rate was set to 1% for protein and peptide identifications. First, absolute quantities were determined for those proteins that had corresponding labeled AQUA peptides (Table S6) using the Trans-Proteomic Pipeline (TPP,V4.4.0) (9). We fitted the relationship between copy numbers and integrated peak areas (as determined by the Progenesis software, NonLinear Dynamics) for the three most intense peaks (10) with a calibration curve to estimate absolute quantities for additional proteins. Extensive validation of this approach had revealed such data to be accurate

within a 2-3fold range (11). Data for three biologically independent samples revealed good reproducibility (Table S3). Enzyme abundance was combined with reported turnover numbers for respective *Salmonella* enzymes (or closely related *E. coli* orthologs) to obtain maximal feasible reaction rates (Table S3). Data were visualized using the pathway tools package (12).

Nutrient utilization assay. We modified a widely used redox assay (13) which measures nutrient utilization based on transfer of reducing equivalents to tetrazolium salts. We grew human clinical isolates of *Salmonella enterica* serovars Typhi and Paratyphi for 24 h as a lawn on LB plates. The bacteria were scraped from the plates and resuspended at a density of 10^9 CFU ml^{-1} in buffer (100 mM NaCl, 30 mM MES pH 6.0, 2 mM NaH_2PO_4, 0.1 mM $MgCl_2$ 1 mM KCl, 0.01% tetrazolium violet). The suspension was distributed on 96well plates containing individual substrates. We used all substrates at comparatively low concentration (0.5 mM, which is 20 to 100fold lower than commonly used in similar redox assays) to more closely mimic relevant in vivo conditions. The plates were incubated at 37°C for 14 h and reduction of the tetrazolium salt to violet formazan was measured photometrically. Background-corrected data were represented in arbitrary formazan units.

Macrophage-like cell culture infection. Raw 264.7 macrophage-like cells were cultured in DMEM cell culture medium containing 10% serum and 0.5 g l^{-1} glucose. Cells were infected with *Salmonella* from stationary cultures at a multiplicity of infection of 30 for 30 min with an initial 5 min. 1100xg centrifugation step. Medium was exchanged against DMEM containing 0.5 g l^{-}

^{1} glucose and 50 mg l^{-1} gentamicin. At 4 hours post infection, medium was exchanged with DMEM containing 0.5 g l^{-1} glucose or 1 g l^{-1} glucose or 0.5 g l^{-1} glucose and 0.5 g l^{-1} mannitol. Cells were washed and lysed 10 h after infection, and aliquots were plated to determine CFU numbers. In some experiments 30 µM dynasore or 30 µM EIPA was added at 4 h post infection.

Confocal microscopy of in vivo dextran endocytosis. Mice were infected with *Salmonella* expressing DsRed. Four days post infection, 100 µg Dextran-AL647 (Invitrogen) was intravenously administered. Two hours later, mice were sacrificed and infected spleen was fixed for 6 h with 4% paraformaldehyde followed by washing in increasing sucrose concentrations. After overnight incubation in 40% sucrose, the tissue was embedded in OCT medium and frozen on dry ice. Cryostat sections were examined with a SP5 confocal microscope (Leica). Images were deconvolved using the "classic maximum likelihood estimation algorithm (CMLE)" of Huygens Remote Manager v2.0.1 (Scientific Volume Imaging) based on metadata of the original images and a theoretical point-spread function. Bitplane Imaris 7.3.0 was used for visualization, rendering, and analysis of the data. After 3d-cropping to focus on the situation around a specific bacterium, we used the surface detection function and the "number of voxels" filter to render *Salmonella* and Dextran surfaces.

Computational modeling of Salmonella metabolism. The consensus genome-scale metabolism reconstruction STMv1 (14) was updated to STMv1.1 (available online) based on recent literature (Table S4). For in vivo

modeling, we modified biomass requirements based on mutant phenotypes (Table S4). We generated an in vivo model using Flux-Balance Analysis (FBA) with the COBRA toolbox (15). Nutrient uptake rates were adjusted to yield consistent results with experimental in vivo growth rates of 26 informative mutants (Supplementary Tables 1, 2 online). The average of 14 amino acid fluxes were used to estimate Asn, Asp, Glu, Gly, Ser uptake (for which we lacked informative mutant data).

For *E. coli* in vitro cultures, maintenance costs of 280'000 ATP molecules s^{-1} have been determined (16). For *Salmonella* in infected host tissues, Flux-balance Analysis of mutant phenotypes suggested elevated maintenance costs of at least 360'000 ATP molecules s^{-1} (see Fig. S4). These higher costs could reflect additional activities (i.e., virulence factor secretion, detoxification of reactive oxygen and nitrogen species, etc.) that *Salmonella* needed to survive in hostile host environments.

We predicted flux distribution using the "minimal total flux" option in the optimize() function. We also determined flux variability in alternative solutions with equal growth rate (0.167 h^{-1}, equivalent to the experimentally determined generation time of 6 h (5)) using the fluxVariability() function. We predicted growth phenotypes for all single gene deletions using the deleteModelGenes() function. Analogous models were built for several isolates of *S*. Typhi, *S*. Paratyphi, *S*. Paratyphi C, and non-typhoidal *Salmonella* (NTS) according to available genome sequences (17-21). The model for *S*. Typhi strain Ty2 is available online (STY2.xml).

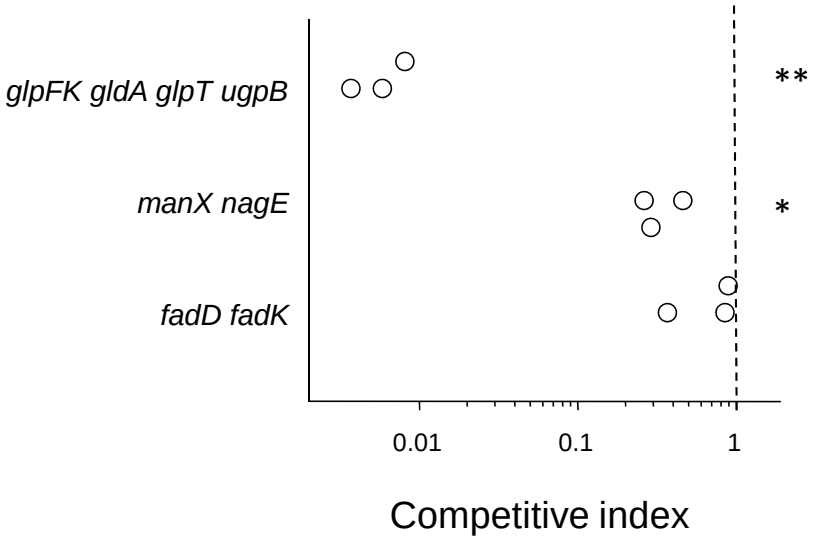
We validated in silico mutant phenotype predictions with experimental data. For this, we classified mutant phenotypes as essential (lethal dose

1000fold higher than wildtype or CI below 0.01, equivalent to a growth of less than 60% based on lethal *Salmonella* loads of 10^7 CFU per spleen), non-detectable (equivalent to a growth rate of 98% or more), or contributing (in between the other two categories). We also used large-scale mutant phenotypes from two recent studies (22, 23). In these cases, we converted the mutant phenotype scores to growth rates and estimated confidence intervals based on the data provided (their Table S3 (22); their Table S3 (23)). In cases, where conflicting data had been reported, we preferentially used data from studies with low infection dose. We disregarded all inconclusive experimental phenotypes for validation of model predictions.

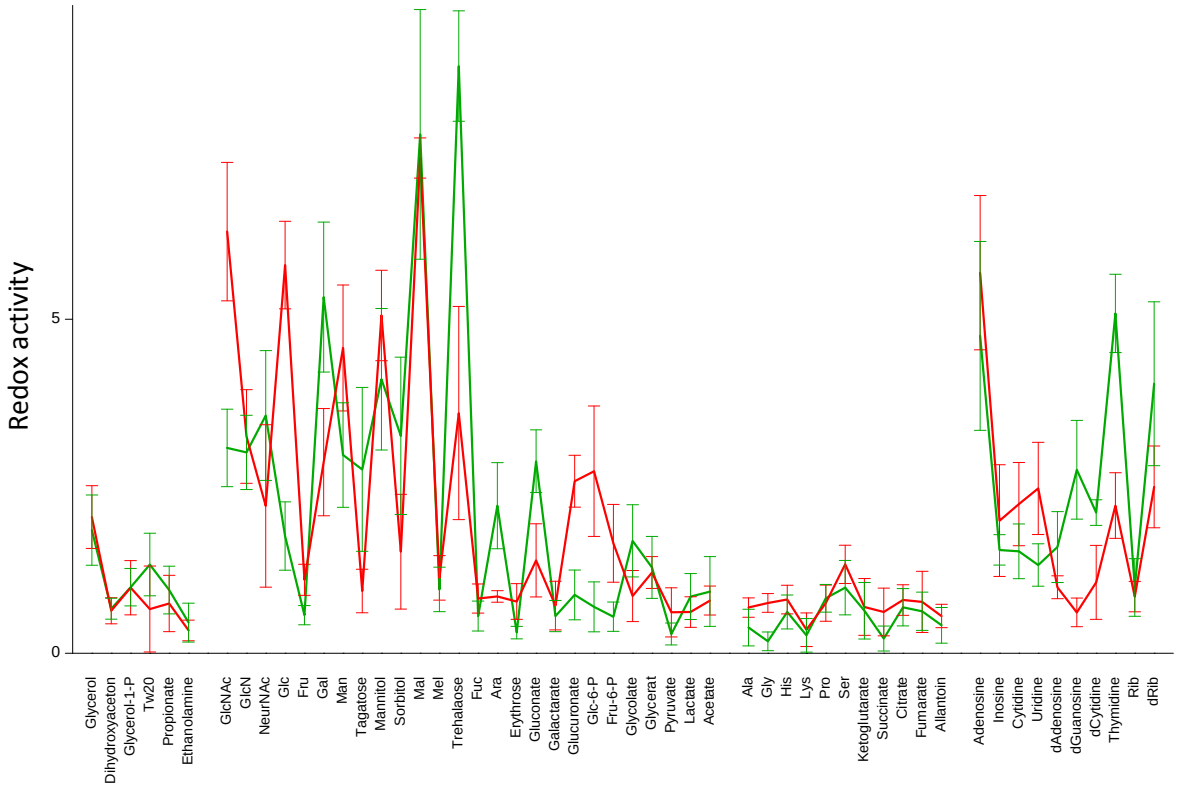
Genome comparisons. Metabolic Pathway predictions for 909 genomes were generated by the MetaCyc consortium (24) and kindly by provided Tomer Altman and Peter Karp on November 22, 2010. We identified 287 mammalian pathogens and 367 environmental organisms in this data set. We merged multiple strains belonging to the same species resulting in 153 pathogen species and 316 environmental species. We then determined how many organisms in each group were capable to degrade a specific nutrient, or to synthesize a certain metabolite.

1. Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640-6645.
2. Thierauf A, Perez G, & Maloy AS (2009) Generalized transduction. *Methods Mol Biol* 501:267-286.
3. Yarmolinsky MB, Wiesmeyer H, Kalckar HM, & Jordan E (1959) Hereditary Defects in Galactose Metabolism in Escherichia Coli Mutants, Ii. Galactose-Induced Sensitivity. *Proc Natl Acad Sci U S A* 45:1786-1791.
4. Ferenci T & Kornberg HL (1973) The utilization of fructose by Escherichia coli. Properties of a mutant defective in fructose 1-phosphate kinase activity. *Biochem J* 132:341-347.
5. Becker D*, et al.* (2006) Robust Salmonella metabolism limits possibilities for new antimicrobials. *Nature.* 440:303-307.
6. Benjamini Y, Drai D, Elmer G, Kafkafi N, & Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125:279-284.
7. Rollenhagen C, Sorensen M, Rizos K, Hurvitz R, & Bumann D (2004) Antigen selection based on expression levels during infection facilitates vaccine development for an intracellular pathogen. *Proc.Natl.Acad.Sci.U.S.A* 101:8739-8744.
8. Maier T*, et al.* (2011) Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol Syst Biol* 7:511.
9. Keller A & Shteynberg D (2011) Software pipeline and data analysis for MS/MS proteomics: the trans-proteomic pipeline. *Methods Mol Biol* 694:169-189.
10. Silva JC, Gorenstein MV, Li GZ, Vissers JP, & Geromanos SJ (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics* 5:144-156.
11. Malmstrom J*, et al.* (2009) Proteome-wide cellular protein concentrations of the human pathogen Leptospira interrogans. *Nature* 460:762-765.
12. Keseler IM*, et al.* (2011) EcoCyc: a comprehensive database of Escherichia coli biology. *Nucleic Acids Res* 39:D583-590.
13. Bochner BR (2009) Global phenotypic characterization of bacteria. *FEMS Microbiol Rev* 33:191-205.
14. Thiele I*, et al.* (2011) A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2. *BMC Syst Biol* 5:8.
15. Becker SA*, et al.* (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* 2:727-738.
16. Feist AM*, et al.* (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.
17. Deng W*, et al.* (2003) Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18. *J.Bacteriol.* 185:2330-2337.
18. Holt KE*, et al.* (2008) High-throughput sequencing provides insights into genome variation and evolution in Salmonella Typhi. *Nat Genet* 40:987-993.

19. Holt KE*, et al.* (2009) Pseudogene accumulation in the evolutionary histories of Salmonella enterica serovars Paratyphi A and Typhi. *BMC Genomics* 10:36.
20. Liu WQ*, et al.* (2009) Salmonella paratyphi C: genetic divergence from Salmonella choleraesuis and pathogenic convergence with Salmonella typhi. *PLoS One* 4:e4510.
21. Kingsley RA*, et al.* (2009) Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res* 19:2279-2287.
22. Santiviago CA*, et al.* (2009) Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLoS Pathog* 5:e1000477.
23. Chaudhuri RR*, et al.* (2009) Comprehensive identification of Salmonella enterica serovar typhimurium genes required for infection of BALB/c mice. *PLoS Pathog* 5:e1000529.
24. Caspi R*, et al.* (2009) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 38:D473-479.

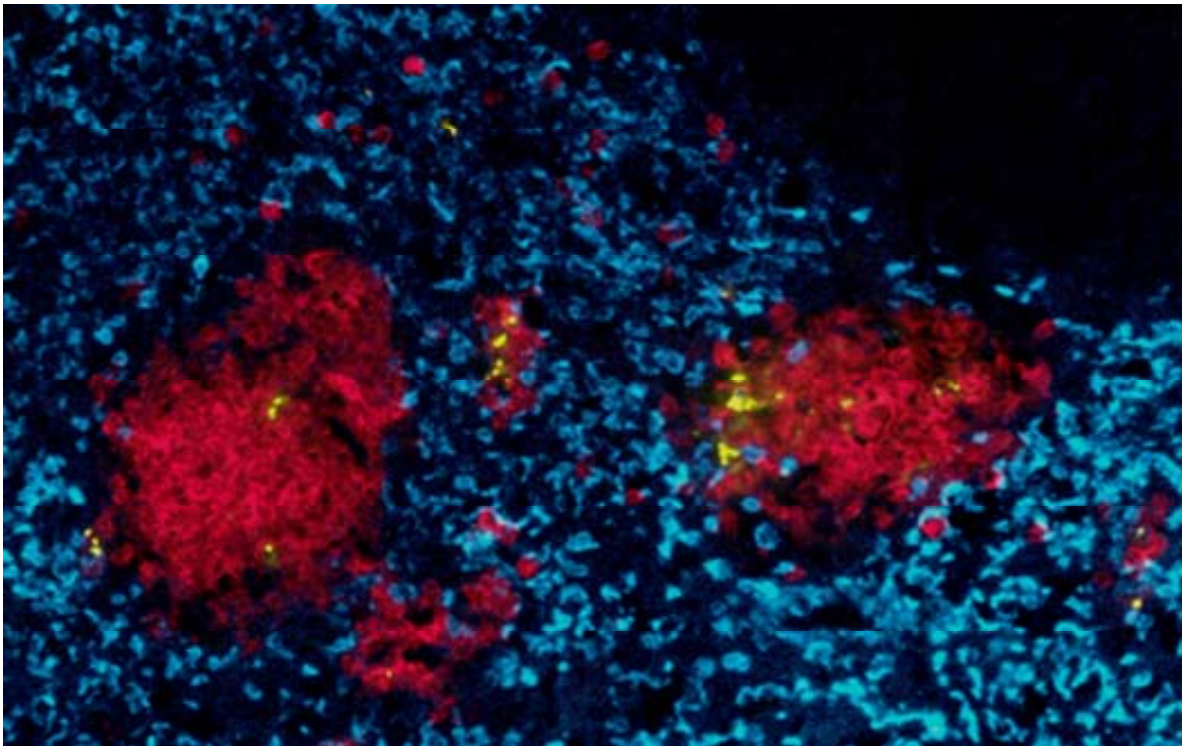


Supporting Figure 1: *Salmonella* mutant phenotypes in competitive infections in genetically resistant 129/Sv mice. Data are represented as competitive indices as compared to wildtype *Salmonella*. A value of 1 indicated identical growth rates of mutant and wildtype. Significance of attenuation was tested with t-test on log-transformed data (*, $P< 0.05$; **, $P< 0.01$).

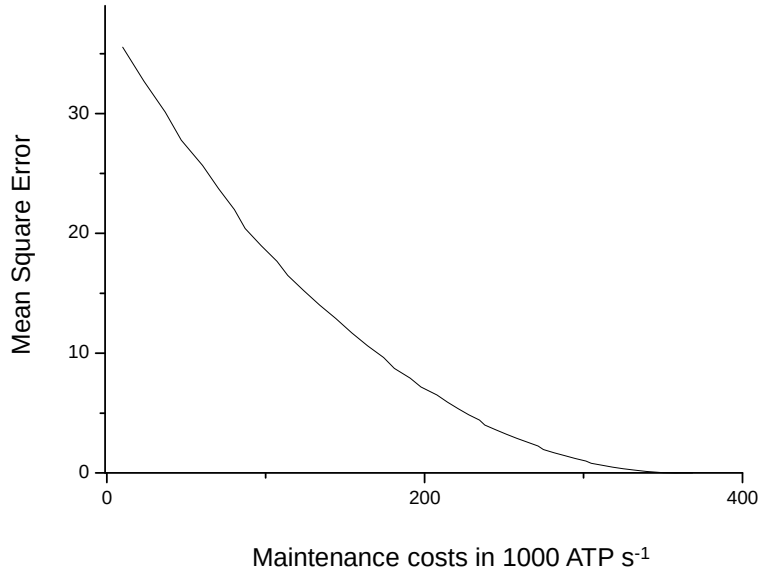


Supporting Figure 2: In vitro nutrient utilization of *S*. Typhi (red) and *S.* Paratyphi (green) human clinical isolates as measured by a redox assay. Data represent averages and SEM's of 5 (*S*. Typhi) or 6 (*S*. Paratyphi) isolates.

***S*. Typhi strain**	**phage type**	**origin, country**
1996-1586	n.c.	Bad Düben, Saxony, carrier since 1966
06-08739	O	Hof, Bavaria, asylum seeker from India
06-05533	46	Berlin, fever, no travel abroad
		Frankfurt/Main,travel to the island
06-02205	E1a	Goa,India
05-07400	C1	Schwerin, carrier
***S*. Paratyphi A**		
5747/2006	untypable	Moers, holiday in Pakistan
06-00724	13	Heilbronn, travel to India
06-02243	1	Idar-Oberstein, travel to India
07-05957	2	Gottmadingen, travel to Pakistan
07-02535	1	Potsdam, travel to India
07-02612	6	Freiburg, travel to India



Supporting Figure 3: Inflammation in *Salmonella*-infected spleen. A spleen cryosection was stained with antibodies to *Salmonella*-LPS (yellow), Ly6C (neutrophils, red), and Ter-119 (erythrocytes, blue). Similar observations were made for 5 mice from three independent experiments.



Supporting Figure 4: Computational analysis of various maintenance costs for *Salmonella* in infected spleen. Mean square errors for predicted vs. experimental mutant growth defects are shown. Experimental data were compatible with elevated in vivo maintenance costs compared to estimates for in vitro cultures (280'000 ATP molecules s^{-1}).

2.3 A small metabolism subnetwork is sufficient for *Salmonella* systemic virulence

(Manuscript in preparation)

Steeb B, Bumann D

Summary:

Experimental data show that the vast majority of *Salmonella* genes have an undetectable impact on virulence. This could reflect limited sensitivity to detect fitness defects, or true massive dispensability.

Here, we used a *Salmonella in vivo* metabolism model in combination with *in vivo* experiments to analyze the virulence contribution of metabolic genes with non-detectable phenotype. Our results indicate that the qualitatively nutrient rich *in vivo* environment makes a majority of genes truly dispensable. In addition, extrinsic and intrinsic redundancy also contributes to gene dispensability. Together, these data suggest that the majority of metabolic enzymes is truly dispensable for replication *in vivo*, suggesting an enormous potential for genome reduction during future niche-adaptation.

Statement of my work:

I have performed all experiments and analyses to generate the data presented in this manuscript. I participated in writing the manuscript.

A small metabolism subnetwork is sufficient for *Salmonella* systemic virulence

Benjamin Steeb[1], Dirk Bumann[1§]

[1]Focal Area Infection Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

[§]Corresponding author

Email address:

dirk.bumann@unibas.ch

Abstract

Background

A vast majority of genes in various pathogens are non-essential for growth in infected host tissues, and this poses severe challenges for development of urgently needed new infection control strategies. To investigate potential causes of gene non-essentiality, we analyzed an extensively validated computational model of *Salmonella* metabolism in infected host tissues.

Results

In silico analyses revealed that a large majority of *Salmonella* enzymes was truly dispensable for in vivo growth. In fact, a minimal metabolism subnetwork of just 364 to 373 enzymes was found to be sufficient to support full *Salmonella* virulence during systemic infection. One important cause for large-scale enzyme dispensability was availability of diverse metabolites in host tissues suggesting that these represented a particularly supportive environment for *Salmonella* growth. In addition, functional redundancy among various *Salmonella* enzymes buffered some metabolic defects. Experimental validation in the mouse typhoid fever model confirmed some synthetically lethal enzyme pairs that might represent suitable combination targets for antimicrobial chemotherapy.

Conclusion

This study revealed that a small metabolic network is sufficient for *Salmonella* virulence. Hundreds of enzymes were dispensable because host environment provided a large variety of nutrients. Some essential metabolic activities could be mediated by two redundant enzymes / pathways suggesting opportunities for novel combination chemotherapies.

Keywords:

Dispensability / flux balance analysis / Metabolism / Pathogen / *Salmonella* / Virulence

Background

Infectious diseases are a major threat to human health worldwide [1]. This situation is worsening because of rapidly increasing antimicrobial resistance of major human pathogens, and this generates an urgent need for new antimicrobials. Unfortunately, the number of suitable targets is severely limited since the vast majority of genes is non-essential for virulence in most pathogens [2-11]. The predominance of genes with weak virulence phenotypes could reflect true gene dispensability, functional redundancy among different genes that collectively mediate important functions, or simply the inability to detect fitness defects below some 5% even when sensitive competitive infection methods are used [2]. Importantly, genes with individually small fitness might collectively still have a substantial impact on fitness (the “marginal benefit” hypothesis: [12]). Causes of gene “dispensability” have been investigated in microbial in vitro cultures in artificial laboratory conditions [13-16], but studies in more relevant natural conditions are rare.

Here, we analyzed the contribution of 1277 metabolic enzymes to *Salmonella* virulence in a mouse typhoid fever model using our recent quantitative computational in vivo metabolism model (Steeb *et al*, unpublished observation). The vast majority of enzymes was found to be truly dispensable, in part because host microenvironments provided a diverse range of metabolites. In contrast, internal *Salmonella* redundancies had a minor impact. Interestingly, metabolism subnetworks of just 364 to 373 enzymes were predicted to be sufficient to support full *Salmonella* virulence during systemic infection, and these

subnetworks could represent potential metabolism end points for the ongoing genome reduction in host-adapted *Salmonella* serovars [17-20]. Together, our data confirmed important challenges for developing new antimicrobials. On the other hand, experimental validation of predicted synthetically lethal pairs revealed several combination targets that could help to develop new antimicrobials to control *Salmonella* infection.

Results

Most *Salmonella* genes contribute weakly or not at all to virulence

The mouse *Salmonella* typhoid fever model is one of the best-characterized in vivo infection models. In particular, virulence phenotypes for thousands of *Salmonella* mutants have been determined. These data suggest that only a small fraction of *Salmonella* genes is essential for virulence, another minor fraction contributes to virulence, while the vast majority of genes have an undetectable impact on virulence. This is evident from reported individual mutant phenotypes (Fig. 1A) as well as two independent unbiased high-throughput studies [3, 4] (Fig. 1B, 1C). *Salmonella* systemic virulence thus represented a typical example for mostly weak mutant phenotypes as observed in many important pathogens.

Metabolism as a suitable subsystem for investigating *Salmonella* gene dispensability

Experimental analysis of potential causes for undetectable phenotypes is challenging. To determine redundancy, *Salmonella* strains with multiple mutations are needed, but construction and in vivo testing of millions of combinations is unfeasible. To detect the collective impact of many genes with undetectable phenotype, one would need *Salmonella* mutants in which all these genes have been inactivated together. Both experimental approaches would require prohibitively large resources.

However, metabolism represents a suitable subsystem that offers unique opportunities for a combination of experimental and rapid computational genome-scale analysis. Importantly, metabolic mutants revealed the same overall distribution of few essential genes and a vast majority of genes with undetectable impact on virulence (see fig. 2A). Moreover, these metabolic phenotypes are accurately predicted by an extensively validated genome-scale computational model of *Salmonella* in vivo metabolism (STMv1.1; Steeb *et al*, unpublished observation). Metabolism thus represented a suitable subsystem to analyze potential causes for mostly undetectable *Salmonella* mutant virulence phenotypes.

The *Salmonella* metabolic network is largely inactive during infection

We used Flux-Balance-Analysis [21, 22] of the computational *Salmonella* in vivo metabolism model STMv1.1 to analyze 1045 *Salmonella* metabolic enzymes with predicted in vivo growth defects below 5% (Fig. 2A, white and light grey). Interestingly, only a minor fraction (68 enzymes; 5% of the total of 1277 enzymes covered in the model) had predicted small in vivo growth defects below the experimental detection threshold (Fig 2A, light grey). These data suggested that limited experimental sensitivity was probably not a major cause of undetectable phenotypes. Instead, the vast majority of *Salmonella* enzymes seemed to be truly dispensable during systemic infection (977 enzymes; 77%).

Large-scale dispensability of metabolic enzymes has previously been observed in other microbial organisms under various in vitro culture conditions. In yeast, most enzymes were dispensable because the metabolic reactions that they could catalyze were inactive in the specific environmental condition [13, 14]. Flux-Balance-Analysis of the *Salmonella* in vivo metabolism model revealed similar widespread reaction inactivity. Specifically, 751 *Salmonella* enzymes catalyzed metabolic reactions that were inactive in the FBA solution with minimal total flux. As expected, removal of all these genes from STMv1.1 resulted in a dramatically reduced model that still supported full *Salmonella* in vivo growth.

These computational predictions were supported by experimental data on *Salmonella* enzyme expression during infection (Steeb *et al*, unpublished observation). In particular, a substantial fraction of enzymes catalyzing metabolic reactions with predicted activity were expressed in vivo at detectable levels (Fig. 4A). This was particularly true for enzymes catalyzing reactions with moderate to high flux rates that were required for *Salmonella* in vivo growth, while we failed to detect several enzymes involved in reactions with low predicted fluxes (Fig. 4B). Small amounts of enzymes might have been sufficient for reactions with low rates, and these were possibly below our experimental in vivo detection limit of a few hundred copies per *Salmonella* cell. Similar correlations between predicted reaction rates and experimental detection of associate enzymes have previously been observed in *E.coli* in vitro cultures [23].

Interestingly, we missed almost all dispensable enzymes that were involved in reactions with predicted fluxes. The majority of these enzymes (61%) had isozymes catalyzing the same reactions ("intrinsic redundancy", see below), and flux-balance analysis could not predict which isozyme was actually used. Other dispensable enzymes catalyzed reactions that could be replaced by equivalent alternative states of the metabolic network as observed in other systems [24, 25]. The flux solution that we used might thus differ somewhat from the actual flux distribution with respect to involvement of dispensable enzymes.

We also detected some enzymes with no apparent role during infection. This could reflect suboptimal metabolic regulation (e.g., expression of enzymes involved in histidine biosynthesis despite pathway blockage by a dysfunctional HisG46 allele [26]) and/or model limitations including an FBA objective function that does not include growth-unrelated aims such as detoxification of reactive oxygen species (e.g., expression of SodCI, AhpC, KatG, etc.).

Together, these computational and experimental data suggested that a majority of *Salmonella* metabolic enzymes was not expressed and truly dispensable for virulence because the reactions that they could catalyze were inactive during infection.

Diverse in vivo nutrition limits requirements for metabolic activity

During infection of mouse spleen, *Salmonella* has access to a wide range of chemically diverse host metabolites, which buffer multiple *Salmonella* metabolic

defects (Steeb *et al*, unpublished observation). To comprehensively analyze the impact of this qualitatively rich nutrition on *Salmonella* activities, we simulated *Salmonella* metabolism in 348 different minimal environments providing various carbon/energy sources and specific supplementation of biomass components. In all of these simulated environments, a different set of *Salmonella* metabolic reactions were predicted to be active compared to qualitatively nutrient rich in vivo conditions. In total, 653 metabolic reactions with no in vivo activity involving 365 genes were conditionally active in one or more simulated minimal environments (blue in fig. 2B and fig. 3). This included biosynthesis pathways for biomass components that were available in infected host tissues but not in simulated media, as well as degradation pathways for nutrients with little in vivo relevance. Qualitatively rich in vivo nutrition thus partially explained why the *Salmonella* metabolic network was largely inactive during infection.

In addition, 646 metabolic reactions remained inactive in all simulated minimal environments (386 dispensable genes, green in fig. 2B and fig. 3). This was in part due to dispensability of certain biomass components such as cobalamine or glycogen during systemic infection, as well as model limitations for predicting alternative pathways with differentially energy costs (e.g., high-affinity vs. low-affinity transport of metabolites across membranes [27]), dead-end pathways because of knowledge gaps, and biomass-unrelated reactions such as detoxification of reactive oxygen species or damaged metabolite salvage (646 reactions; Fig. 3 green).

Functional redundancy causes enzyme dispensability in active reactions

A subset of 554 reactions associated with 526 enzymes was predicted to be active in vivo. This reaction network was somewhat larger compared to previously studied in vitro cultures of other microbes that were based on less comprehensive metabolic reconstructions [28]. A surprisingly large fraction of these reactions (210 reactions, 38%) was associated with dispensable enzymes (226 enzymes). The majority of these reactions was actually required for growth but an involved enzyme could be replaced by one or several alternative isozymes (”intrinsic redundancy” [29]; 143 reactions associated with 138 enzymes; see fig. 3, red). Intrinsic redundancy was particularly frequent in central carbon metabolism as observed previously [30]. In addition, some reactions were active but dispensable because they could be replaced by some other reaction / overlapping pathways (“extrinsic redundancy” [29]; 44 reactions associated with 69 dispensable genes; Fig. 3, yellow).

Mutually compensating enzymes could provide largely unexplored opportunities for antimicrobial combination chemotherapy. Specifically, simultaneous inhibition of redundant enzymes could result in *Salmonella* growth arrest. To comprehensively analyze such opportunities, we predicted all possible 616’605 double gene deletion phenotypes. The data revealed 88 dispensable genes that formed synthetically lethal pairs with less than 60% growth compared to wildtype. This set of synthetic pairs partially overlapped with previous predictions [30, 31].

To experimentally validate some of the predicted pairs, we constructed respective *Salmonella* double mutants and tested them in the mouse typhoid fever model. Among 13 tested mutants, six were indeed avirulent (competitive index below 0.01 in mixed infections with wildtype *Salmonella*; such slow mutant growth can be successfully controlled by the host [2] (see fig. 5 and additional table 1).

On the other hand, several predicted synthetically lethal mutations did not abrogate *Salmonella* virulence. These discrepancies suggested knowledge gaps in the computational model. In particular, auxotrophic mutants that were deficient for all known supplement uptake systems were still able to grow in supplemented minimal media and in infected mice, indicating additional yet uncharacterized transport capabilities. Indeed, the genome of *Salmonella* contains hundreds of transporters with no experimental data, and prediction of putative substrates is difficult. We attempted to assign possible transporters for methionine, chloride, thiamin, pyridoxal, and histidine based on sequence similarities to known transporters and/or indirect evidence. However, additional inactivation of 8 putative transporters in five *Salmonella* mutants still did not abrogate growth in supplemented minimal media indicating yet other transport mechanisms (see additional table 1).

Taken together, these data reveal interesting opportunities for antimicrobial combination therapy, but also highlight important issues for future experimentation and improvement of metabolic reconstructions.

Small non-redundant minimal metabolic networks might be sufficient for *Salmonella* in vivo growth

Inactivity and functional redundancy caused large-scale dispensability of *Salmonella* metabolic enzymes. These data suggested that a small set of *Salmonella* enzymes might be sufficient for full in vivo growth. We investigated this issue by sequential random removal of genes until the predicted growth rate decreased. The results revealed strongly reduced minimal metabolic networks with only 364 to 373 (28% of the original network) that would still support normal *Salmonella* growth in infected tissues. Interestingly, the various independent genome reduction trajectories converged on highly similar networks with 300 invariantly present enzymes suggesting a dominant, largely non-redundant core of commonly required metabolic capabilities (Fig 3, black). Similar small genomes have been predicted for obligate intracellular symbionts such as *Buchnera* or *Wigglesworthia* [32] suggesting comparable in vivo conditions for intracellular pathogens and symbionts. Together these data again support true dispensability of a vast majority of *Salmonella* metabolic enzymes during systemic infection.

Discussion

Experimental analysis of diverse pathogens in various infection models has revealed generally small numbers of essential virulence genes and vast majorities of genes with undetectable impact on virulence. This could just reflect technical limitations to detect small virulence contributions. Alternatively, many of these genes might be truly dispensable for virulence.

In this study, we investigated *Salmonella* metabolic enzymes that have no detectable impact on *Salmonella* virulence in a well-characterized mouse typhoid fever model. Analysis of an extensively validated genome-scale computational model revealed that only few metabolic enzymes had minor fitness contributions that might escape experimental detection. In contrast, the vast majority of *Salmonella* enzymes was truly dispensable. The main cause of dispensability was reaction inactivity due to external supplementation with diverse nutrients as previously observed for yeast in vitro cultures [13, 14]. In addition, certain biomass components were not required during infection, and redundant isozymes and alternative pathways buffered some enzyme inactivation.

Together these data suggested that hundreds of *Salmonella* genes were truly dispensable in a mouse typhoid fever model. Experimental validation of this hypothesis will require extensive genome reduction through hundreds of gene deletions. This is principally feasible [33-35] but would require prohibitive resources. On the other hand, evolution of various pathogens including *Salmonella* is associated with striking genome reduction by gene inactivation through frameshift mutations / nonsense mutations or gene loss (*Salmonella*

Typhi [17, 18], *Mycobacterium leprae* [36], *Buchnera aphidicola* [37], *Rickettsia prowazekii* [38], *Mycoplasma pneumonia* [39]), indicating minimal metabolic requirements for pathogens residing in specific stable host niches. The small subnetwork identified in this study might represent a potential endpoint for host-adapted *Salmonella* serovars such as *S.* Typhi and *S.* Paratyphi that cause mainly systemic disease. Indeed, partial genome degradation of these serovars has already inactivated 81 enzymes [20], which were involved in *Salmonella* metabolic activities predicted to be dispensable during systemic infection.

On the other hand, most of the genes that were found to be dispensable in the mouse typhoid fever model are still broadly conserved among diverging *Salmonella* lineages suggesting some relevant fitness benefits. It is important to note that we investigated just one ecological niche - the infected mouse spleen - among the various habitats that are relevant in the *Salmonella* infection / transmission cycle. In vivo “dispensable” yet conserved genes likely serve relevant functions in some of these other habitats [40, 41]. As examples, tetrathionate reduction and ethanolamine utilization are dispensable during systemic infection, but highly relevant in the intestinal lumen [42, 43]. Finally, *Salmonella enterica* serovar Typhimurium has a broad host range and several genes might be specifically required in important host species other than mice [44-46].

The small number of essential genes in pathogens severely limits the target space for novel antimicrobials. Combinatorial inhibition of redundant enzymes could offer additional, yet largely unexplored opportunities. Some

isozyme pairs that catalyze essential reactions might be particularly promising as it might be possible to identify promiscuous inhibitors that block both isozymes. The alanine racemases involved in biosynthesis of the essential cell wall component D-alanine could represent a classical example. Indeed, a *Salmonella* double mutant *alr dadX* was avirulent although weak side reactions of other enzymes obviously still provided small amounts of D-alanine to enable residual in vitro and in vivo growth as previously observed for *E. coli* [47]. Such cryptic metabolic activities provide critical opportunities for rapid resistance development, which could make this particular target combination appear less attractive.

In contrast to isozymes, synthetically lethal combinations of unrelated enzymes will probably require two different inhibitors for successful combination therapy. This might potentiate drug development problems as two different compounds would require appropriate pharmacodynamics and pharmacokinetics to simultaneously act on the same pathogen cells in infected tissue. However, the clinically highly successful example of β-lactam antibiotics / β-lactamase inhibitor combinations demonstrate that such combination therapy against individually dispensable targets is possible [48]. Some of the synthetically lethal pairs that we could confirm in experimental mouse infections in this study might thus represent suitable start points for development of urgently needed new antimicrobial chemotherapies.

Conclusion:

In this manuscript, we identified multiple mechanisms that cause extensive gene dispensability in the pathogen *Salmonella* Typhimurium during infection. The fact that a small metabolic subnetwork was sufficient for virulence severely limited the target space for novel antimicrobials in this pathogen. However, some of the synthetically lethal combinations that were identified and validated in this study could represent suitable targets for novel antimicrobial chemotherapies.

Methods

Bacterial genetics. We used strain *Salmonella enterica* serovar Typhimurium SL1344 *hisG xyl* [26] as parental wild type strain. *Salmonella* mutants were constructed by lambda red-recombinase mediated allelic replacement [49] followed by general transduction using phage P22 *int* [50]. In some cases, resistance cassettes flanked with FRT sites were removed using FLP recombinase [49]. Strains were cultivated on Lennox LB medium containing 90 µg ml-1 streptomycin, 50 µg ml-1 kanamycin, 20 µg ml-1 chloramphenicol, and/or 100 µg ml-1 ampicillin. Auxotrophs were supplemented with 260 µM L-methionine (*yaeC metA*), 15 µM D-arabinose-5-phosphate / 10 µM glucose-6-phosphate (*gutQ yrbH*),15 µM mannose-1-phosphate / 10 µM glucose-6-phosphate (*cpsG rfbK*), 1 µM thiamin (*yabJ thiI*), 1 mM β-pantothenate (*panC panF*), 100 µM nicotinamide mononucleotide (*nadC pncB*), 10 µM pyridoxin (*pdxA STM0163*), 250 µM D-alanine (*alr dadX*). Mutant *folD purT* was supplemented with nucleosides (Thymidine [300 µM], uridine [300 µM], guanosine [300 µM], cytidine [300 µM], adenosine [300 µM] and 2-deoxy-D-ribose [300 µM]), amino acids (RPMI 1640 Amino acids solution (50x, Sigma) [2.5x final concentration], L-alanine [250 µM] and L-glutamine [250 µM]) and vitamins (RPMI 1640 Vitamins Solution (100x, Sigma) [1x final concentration]). Phenotypes of auxotrophic mutants were validated using M9 minimal media containing appropriate supplements.

Mouse infections. For competitive infections, wildtype and mutant *Salmonella* expressing various fluorescent proteins (our unpublished observation) were cultured in LB medium to late log phase. The strains were mixed and a total 1000 - 4000 CFU were injected into the tail veins of female, 8 to 12 weeks old BALB/c mice. Four days post-infection, mice were sacrificed and total bacterial loads in spleen were determined by plating. Individual strains were quantified by multi-color flow cytometry using a Fortezza flow cytometer equipped with 488 nm and 561 nm lasers (BD Biosciences) (our unpublished observation). The data were used to calculate competitive indices compared to wildtype *Salmonella* (CI = output ratio / input ratio). We converted CI values to growth rates based on a 6 h in vivo generation time [2]. Statistical significance was analyzed using one-sample t-test (a parametric test was appropriate based on the normal distribution of growth rates [2]) and determination of false discovery rates [53].

Computational modeling of Salmonella metabolism. We used a genome-scale computational *Salmonella* in vivo metabolism model STMv1.1, an updated version of the consensus genome-scale metabolism reconstruction STMv1 ([31], Steeb *et al*, unpublished observation). The in vivo model was analyzed with Flux-Balance Analysis (FBA) [22] using MatLab and the COBRA toolbox [51]. We predicted flux distribution in vivo and in different minimal media using the "minimal total flux" option in the optimize() function. Genes that did not participate in reactions with flux were scored "inactive". The contribution of metabolic reactions was determined using the changeRxnBounds() function. We predicted growth phenotypes for all single, double and multiple gene deletions using the

deleteModelGenes() and doubleGeneDeletion() functions. To obtain minimal metabolism subnetworks, we randomly selected one gene and deleted it. If the reduced network still permitted normal growth, random gene selection and deletion was continued. We ran this procedure 30 times with independent trajectories. Intrinsic redundancy was identified based on growth-contributing reactions that nevertheless contained dispensable genes. To assess extrinsic redundancy, we constrained the network to the actual minimal total flux solution. This constrain removed any compensation opportunity through normally inactive pathways. Reactions that showed a contribution only in this constrained network, but not in the fully unconstrained network were scored as dispensable through extrinsic redundancy. Data were visualized using the pathway tools package [52].

Classification of virulence categories. We classified in vivo / in silico mutant phenotypes as essential if growth was less than 60% of the wildtype rate (equivalent to a lethal dose 1000fold higher than wildtype, or CI values below 0.01). Non-detectable phenotypes had growth rate phenotypes below the detection limit (95% wildtype growth rate), contributing phenotypes were in between the other two categories.

List of abbreviations used

CI, competitive index; **FBA**, flux balance analysis; **gDW**, gramm dry weight; **wt**, wildtype;

Competing interests

The authors declare that they have no competing interests.

Authors' Contributions

BS performed experiments. DB and BS performed in silico analyses. DB and BS wrote the manuscript.

Acknowledgments

We thank Beatrice Claudi, Alain Mazé and Neil Burton for technical help with in vivo experiments. D. Bumann and co-workers are funded by the Swiss National Science Foundation (31003A-121834), SystemsX (RTD project BattleX) and Deutsche Forschungsgemeinschaft (SFB641-A9, SPP1316 Bu971/6).

References

1. Norrby SR, Nord CE, Finch R: **Lack of development of new antimicrobial drugs: a potential serious threat to public health.** *Lancet Infect Dis* 2005, **5:**115-119.
2. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, Mann M, Bumann D: **Robust Salmonella metabolism limits possibilities for new antimicrobials.** *Nature* 2006, **440:**303-307.
3. Chaudhuri RR, Peters SE, Pleasance SJ, Northen H, Willers C, Paterson GK, Cone DB, Allen AG, Owen PJ, Shalom G, et al: **Comprehensive identification of Salmonella enterica serovar typhimurium genes required for infection of BALB/c mice.** *PLoS Pathog* 2009, **5:**e1000529.
4. Santiviago CA, Reynolds MM, Porwollik S, Choi SH, Long F, Andrews-Polymenis HL, McClelland M: **Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice.** *PLoS Pathog* 2009, **5:**e1000477.
5. Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, Holden DW: **Simultaneous identification of bacterial virulence genes by negative selection.** *Science* 1995, **269:**400-403.
6. Kavermann H, Burns BP, Angermuller K, Odenbreit S, Fischer W, Melchers K, Haas R: **Identification and characterization of Helicobacter pylori genes essential for gastric colonization.** *J Exp Med* 2003, **197:**813-822.
7. Maroncle N, Balestrino D, Rich C, Forestier C: **Identification of Klebsiella pneumoniae genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis.** *Infect Immun* 2002, **70:**4729-4734.
8. Edelstein PH, Edelstein MA, Higa F, Falkow S: **Discovery of virulence genes of Legionella pneumophila by using signature tagged mutagenesis in a guinea pig pneumonia model.** *Proc Natl Acad Sci U S A* 1999, **96:**8190-8195.
9. Sun YH, Bakshi S, Chalmers R, Tang CM: **Functional genomics of Neisseria meningitidis pathogenesis.** *Nat Med* 2000, **6:**1269-1273.
10. Chiang SL, Mekalanos JJ: **Use of signature-tagged transposon mutagenesis to identify Vibrio cholerae genes critical for colonization.** *Mol Microbiol* 1998, **27:**797-805.
11. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C: **Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis.** *Mol Microbiol* 1999, **34:**257-267.
12. Thatcher JW, Shaw JM, Dickinson WJ: **Marginal fitness contributions of nonessential genes in yeast.** *Proc Natl Acad Sci U S A* 1998, **95:**253-257.
13. Papp B, Pal C, Hurst LD: **Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast.** *Nature* 2004, **429:**661-664.
14. Blank LM, Kuepfer L, Sauer U: **Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast.** *Genome Biol* 2005, **6:**R49.
15. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A, et al: **Phenotypic landscape of a bacterial cell.** *Cell* 2011, **144:**143-156.

16. Hillenmeyer ME, Fung E, Wildenhain J, Pierce SE, Hoon S, Lee W, Proctor M, St Onge RP, Tyers M, Koller D, et al: **The chemical genomic portrait of yeast: uncovering a phenotype for all genes.** *Science* 2008, **320:**362-365.
17. Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, et al: **High-throughput sequencing provides insights into genome variation and evolution in Salmonella Typhi.** *Nat Genet* 2008, **40:**987-993.
18. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT, et al: **Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.** *Nature* 2001, **413:**848-852.
19. Liu WQ, Feng Y, Wang Y, Zou QH, Chen F, Guo JT, Peng YH, Jin Y, Li YG, Hu SN, et al: **Salmonella paratyphi C: genetic divergence from Salmonella choleraesuis and pathogenic convergence with Salmonella typhi.** *PLoS One* 2009, **4:**e4510.
20. Holt KE, Thomson NR, Wain J, Langridge GC, Hasan R, Bhutta ZA, Quail MA, Norbertczak H, Walker D, Simmonds M, et al: **Pseudogene accumulation in the evolutionary histories of Salmonella enterica serovars Paratyphi A and Typhi.** *BMC Genomics* 2009, **10:**36.
21. Varma A, Palsson BO: **Predictions for oxygen supply control to enhance population stability of engineered production strains.** *Biotechnol Bioeng* 1994, **43:**275-285.
22. Oberhardt MA, Palsson BO, Papin JA: **Applications of genome-scale metabolic reconstructions.** *Mol Syst Biol* 2009, **5:**320.
23. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, Adkins JN, Schramm G, Purvine SO, Lopez-Ferrer D, et al: **Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models.** *Mol Syst Biol* 2010, **6:**390.
24. Mahadevan R, Schilling CH: **The effects of alternate optimal solutions in constraint-based genome-scale metabolic models.** *Metab Eng* 2003, **5:**264-276.
25. Reed JL, Palsson BO: **Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states.** *Genome Res* 2004, **14:**1797-1805.
26. Hoiseth SK, Stocker BA: **Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines.** *Nature* 1981, **291:**238-239.
27. Ammendola S, Pasquali P, Pistoia C, Petrucci P, Petrarca P, Rotilio G, Battistoni A: **High-affinity Zn2+ uptake system ZnuABC is required for bacterial zinc homeostasis in intracellular environments and contributes to the virulence of Salmonella enterica.** *Infect Immun* 2007, **75:**5867-5876.
28. Nishikawa T, Gulbahce N, Motter AE: **Spontaneous reaction silencing in metabolic optimization.** *PLoS Comput Biol* 2008, **4:**e1000236.
29. Hartman JLt, Garvik B, Hartwell L: **Principles for the buffering of genetic variation.** *Science* 2001, **291:**1001-1004.
30. Suthers PF, Zomorrodi A, Maranas CD: **Genome-scale gene/reaction essentiality and synthetic lethality analysis.** *Mol Syst Biol* 2009, **5:**301.

31. Thiele I, Hyduke DR, Steeb B, Fankam G, Allen DK, Bazzani S, Charusanti P, Chen FC, Fleming RM, Hsiung CA, et al: **A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2.** *BMC Syst Biol* 2011, **5:**8.
32. Pal C, Papp B, Lercher MJ, Csermely P, Oliver SG, Hurst LD: **Chance and necessity in the evolution of minimal metabolic networks.** *Nature* 2006, **440:**667-670.
33. Mizoguchi H, Sawano Y, Kato J, Mori H: **Superpositioning of deletions promotes growth of Escherichia coli with a reduced genome.** *DNA Res* 2008, **15:**277-284.
34. Posfai G, Plunkett G, 3rd, Feher T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M, et al: **Emergent properties of reduced-genome Escherichia coli.** *Science* 2006, **312:**1044-1046.
35. Hashimoto M, Ichimura T, Mizoguchi H, Tanaka K, Fujimitsu K, Keyamura K, Ote T, Yamakawa T, Yamazaki Y, Mori H, et al: **Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome.** *Mol Microbiol* 2005, **55:**137-149.
36. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, et al: **Massive gene decay in the leprosy bacillus.** *Nature* 2001, **409:**1007-1011.
37. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H: **Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.** *Nature* 2000, **407:**81-86.
38. Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC, Podowski RM, Naslund AK, Eriksson AS, Winkler HH, Kurland CG: **The genome sequence of Rickettsia prowazekii and the origin of mitochondria.** *Nature* 1998, **396:**133-140.
39. Yus E, Maier T, Michalodimitrakis K, van Noort V, Yamada T, Chen WH, Wodke JA, Guell M, Martinez S, Bourgeois R, et al: **Impact of genome reduction on bacterial metabolism and its regulation.** *Science* 2009, **326:**1263-1268.
40. Soyer OS, Pfeiffer T: **Evolution under fluctuating environments explains observed robustness in metabolic networks.** *PLoS Comput Biol* 2010, **6**.
41. Semenov AM, Kuprianov AA, van Bruggen AH: **Transfer of enteric pathogens to successive habitats as part of microbial cycles.** *Microb Ecol* 2010, **60:**239-249.
42. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, et al: **Gut inflammation provides a respiratory electron acceptor for Salmonella.** *Nature* 2010, **467:**426-429.
43. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Baumler AJ: **Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota.** *Proc Natl Acad Sci U S A* 2011, **108:**17480-17485.
44. Morgan E, Campbell JD, Rowe SC, Bispham J, Stevens MP, Bowen AJ, Barrow PA, Maskell DJ, Wallis TS: **Identification of host-specific colonization factors**

of Salmonella enterica serovar Typhimurium. *Mol Microbiol* 2004, **54:**994-1010.

45. Richardson EJ, Limaye B, Inamdar H, Datta A, Manjari KS, Pullinger GD, Thomson NR, Joshi RR, Watson M, Stevens MP: **Genome sequences of Salmonella enterica serovar typhimurium, Choleraesuis, Dublin, and Gallinarum strains of well- defined virulence in food-producing animals.** *J Bacteriol* 2011, **193:**3162-3163.
46. Stevens MP, Humphrey TJ, Maskell DJ: **Molecular insights into farm animal and zoonotic Salmonella infections.** *Philos Trans R Soc Lond B Biol Sci* 2009, **364:**2709-2723.
47. Kang L, Shaw AC, Xu D, Xia W, Zhang J, Deng J, Woldike HF, Liu Y, Su J: **Upregulation of MetC is essential for D-alanine-independent growth of an alr/dadX-deficient Escherichia coli strain.** *J Bacteriol* 2011, **193:**1098-1106.
48. Bush K, Macielag MJ: **New beta-lactam antibiotics and beta-lactamase inhibitors.** *Expert Opin Ther Pat* 2010, **20:**1277-1293.
49. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.** *Proc Natl Acad Sci U S A* 2000, **97:**6640-6645.
50. Thierauf A, Perez G, Maloy AS: **Generalized transduction.** *Methods Mol Biol* 2009, **501:**267-286.
51. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.** *Nat Protoc* 2007, **2:**727-738.
52. Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muniz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M, Altman T, et al: **EcoCyc: a comprehensive database of Escherichia coli biology.** *Nucleic Acids Res* 2011, **39:**D583-590.
53. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I: **Controlling the false discovery rate in behavior genetics research.** *Behav Brain Res* 2001, **125:**279-284.

Figure legends

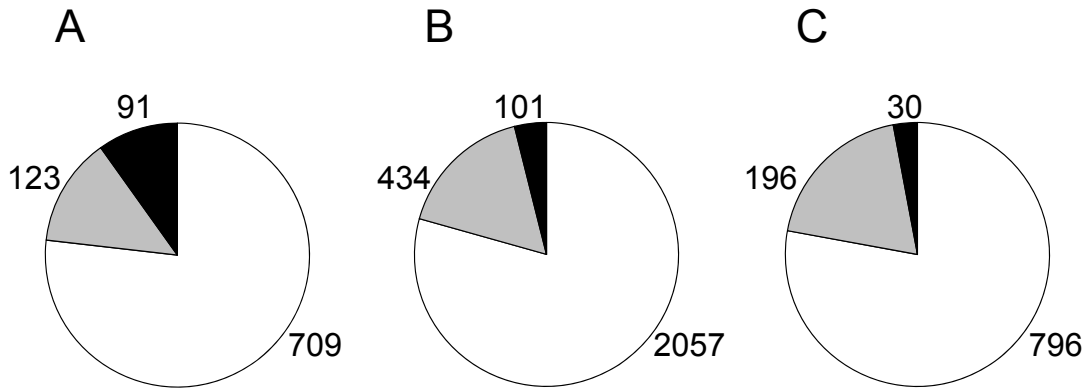


Figure 1: Summary of *Salmonella* virulence phenotypes in the mouse typhoid fever model. Figure 1 displays the proportion of genes essential for virulence **(black)**, genes with a contributing phenotype **(grey)** and genes without a detectable phenotype **(white)** (for category definition, see Methods). Data were obtained from literature describing individually studied mutants (**A**), or two unbiased large-scale screens (**B** [3], **C** [4]). For each group of essential, contributing, and non-detectable phenotypes, the corresponding number of genes is shown.

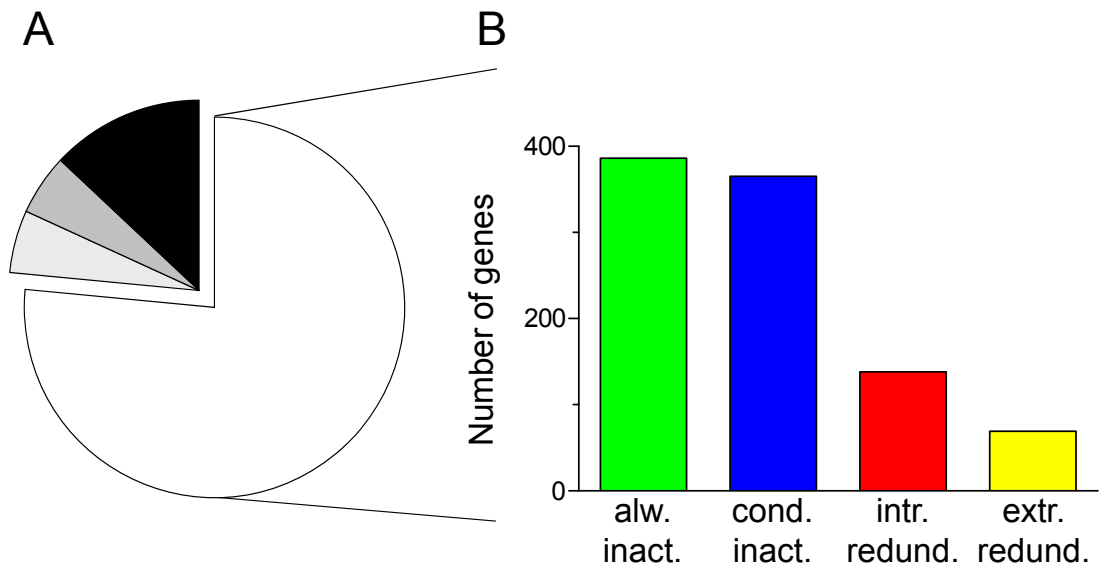


Figure 2: Analysis of genes with non-detectable phenotype that have no contribution to virulence.

(A) In silico analyses based on the model STMv1.1 predict that in vivo 166 genes are essential for virulence **(black)**, 66 are contributing **(dark grey)** and 1045 have a non-detectable phenotype **(light grey and white)**. 977 from the 1045 genes with non-detectable phenotype don't contribute to virulence at all **(white)**. **(B)** The 977 genes with non-detectable phenotype and without virulence contribution can be grouped into categories depending on the reason for their dispensability into always inactive genes **(green)**, conditionally inactive genes **(blue)**, genes dispensable due to intrinsic redundancy **(red)** and genes dispensable due to extrinsic redundancy **(yellow)**.

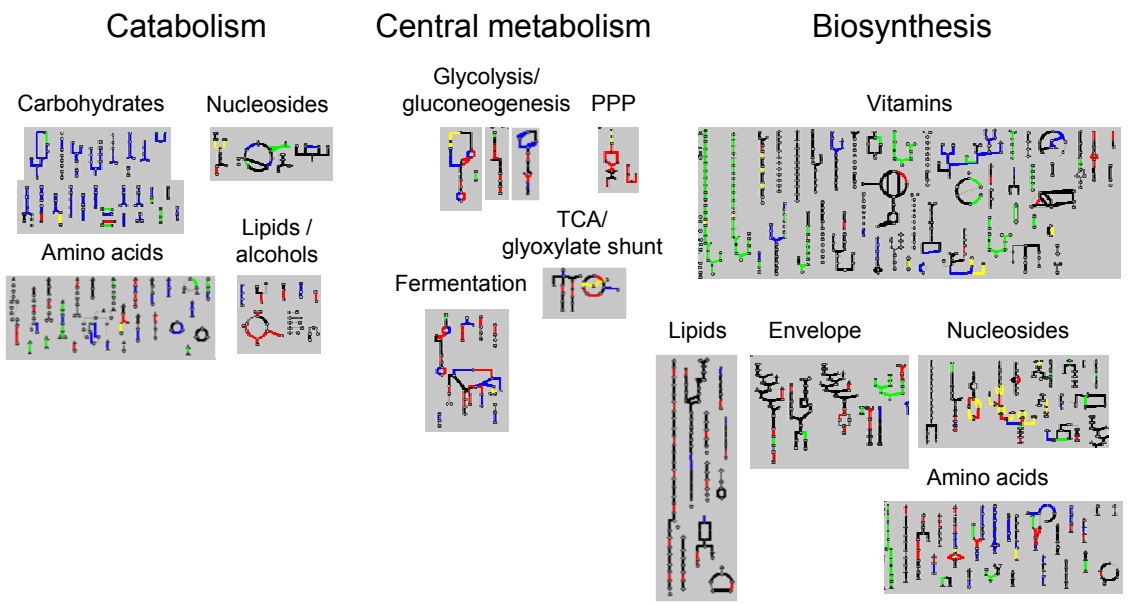


Figure 3: Gene dispensability in *Salmonella* Typhimurium in vivo metabolism.

In this figure, relevant parts of Salmonella metabolism is displayed, grouped into the categories catabolism, central metabolism and biosynthesis. (**Black**): Genes are always part of the minimal metabolic backbone needed for uninhibited in vivo growth (300 genes). (**Red**): Genes are in vivo active, but dispensable due to intrinsic redundancy (138 genes). (**Yellow**): Genes are in vivo active, but dispensable due to extrinsic redundancy (69 genes). (**Blue**): Genes are in vivo inactive, but can become active in another environment (365 genes). (**Green**): Genes that were inactive in all minimal FBA predictions (386 genes).

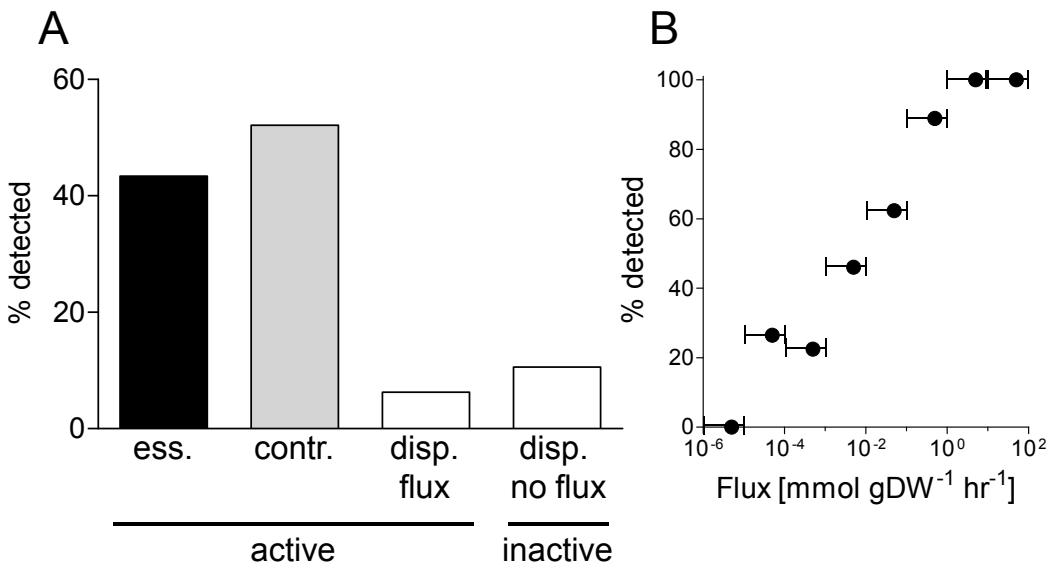


Figure 4: Analysis of ex vivo detected enzymes.

(A) Percentage of detected ex vivo proteins for different phenotype categories. Ex vivo detected enzymes were predicted to be essential (**black**), contributing (**grey**) or dispensable for virulence (**white**) and were associated with reactions with flux (**active**) or without flux (**inactive**). **(B)** Percentage of essential to contributing reactions with at least one detected protein plotted against the catalyzed flux of the reaction (in mmol gDW^{-1} hr^{-1}). Bin size of one log scale is indicated as bars.

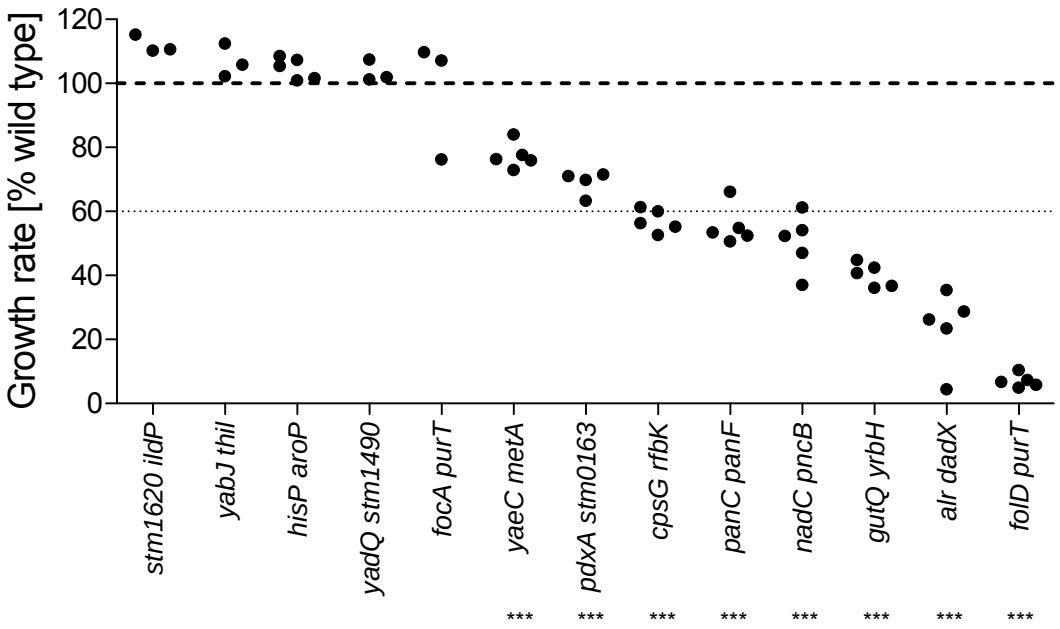


Figure 5: Growth rates of constructed double mutants. Average growth rates of mutants were calculated based on competitive indices to the wildtype, deduced from bacterial loads in spleen. An average wild type division time of 6 h was used for calculation of growth rates [2]. Significance levels are given for false discovery rates [53] based on one-tailed, one-sample t-tests (*: 0.05 < q < 0.01; **: 0.01 < q < 0.001; ***: q < 0.001).

Supplemental table 1

Mutated loci	Gene names	Growth rate in % of wildtype	*P*-value	FDR[1]	Expected effect of double deletion	Potential explanation for unexpected in vivo growth	Alternative transporters tested[2]
stm1620 stm3692	*stm1620 ilvDP*	112.1 ± 2.8	-	-	Accumulation of glycolate	Maybe flux reversion of irreversible reaction upon glycolate accumulation (e.g. export via importer ActP (STM4273) or oxidation via glycolate dehydrogenase (YiaE (STM3646) or YcdW (STM1135))	-
stm0106 stm0425	*yabJ thiI*	106.8 ± 5.2	-	-	Thiamine auxotrophy and transport deficiency	In addition to *yabLKJ* (*thiBPQ*) another unknown thiamine transport system is present in *Salmonella* **[54]**	*yieG* (*stm3851*)
stm2351 stm0150	*hisP aroP*	104.8 ± 3.4	-	-	Histidine uptake deficiency in histidine auxotroph SL1344	The presence of another unknown histidin transport system is indicated by the possibility of *in vitro* histidine supplementation of SL1344 *hisP aroP* (also suggested in **[55]**)	*pheP* (*stm0568*)
stm0203 stm1490	*yadQ stm1490*	103.5 ± 3.4	-	-	Chloride uptake deficiency	Maybe additional, unspecific chloride transporters are existing. The construction of this gene deletion combination was already described in *E.coli* **[56]**	*yfeO* (*stm2404*), *stm1527*
stm0974 stm1883	*focA purT*	97.7 ± 18.6	0.6594	>0.05	Accumulation of formate	Unspecific formate excretion or activity of aerobically inactive formiat hydrogen lyase FdhF (STM4285)	-
stm0245 stm4182	*yaeC metA*	77.4 ± 4.1	<0.0001	<0.001	Methionine auxotrophy and transport deficiency	In addition to *abc-yaeE-yaeC* (*metNIQ*), another unknown or unspecific methionine transport system is present in *Salmonella* **[57]**	*aroP* (*stm0150*), *brnQ* (stm*0399*), *leuE (stm1270*)
stm0091 stm0163	*pdxA stm0163*	68.9 ± 3.8	<0.0001	<0.001	Pyridoxal auxotrophy	*In vivo* pyridoxal supplementation most likely possible via a so far unidentified uptake system for pyridoxamine/ pyridoxal/ pyridoxine reported in literature **[58-60]**	*codB* (*stm3333*), *allP* (*stm0522*)
stm2104 stm2083	*cpsG rfbK*	57.1 ± 3.6	<0.0001	<0.001	Synthesis of full O-antigen blocked	Excess of D-mannose-1-P before infection could be the reason for residual *in vivo* growth	-
stm0181 stm3382	*panC panF*	55.5 ± 6.1	<0.0001	<0.001	Pantothenate auxotrophy and uptake deficiency	Pantothenate supplementation of SL1344 *panC panF* most likely possible via unspecific uptake or a so far unidentified uptake system	-
stm0145 stm1004	*nadC pncB*	50.4 ± 9.0	<0.0001	<0.001	Nicotinic acid / β-NMN auxotrophy	Low amounts of nicotinic acid / β-NMN available *in vivo*	-
stm2838 stm3315	*gutQ yrbH*	40.2 ± 3.7	<0.0001	<0.001	Synthesis for LPS precursor KDO (2-keto-3-deoxy-octonate) blocked; D-arabinose-5-P auxotroph	Excess of D-arabinose 5-P supplementation before infection could be the reason for residual *in vivo* growth. The construction of this gene deletion combination was already described in *E.coli* **[61]**	-
stm4247 stm1802	*alr dadX*	23.6 ± 11.6	<0.0001	<0.001	Peptidoglycan synthesis blocked	Low *in vivo* growth rates are most likely possible through cryptic activity of other enzymes (e.g. MetC **[47]**). The construction of this gene deletion combination was already described in *Salmonella* **[62]**	-
stm0542 stm1883	*folD purT*	7.1 ± 2.1	<0.0001	<0.001	Tetrahydrofolate synthesis blocked	-	-

[1]: False discovery rate

[2]: None with enhanced *in vitro* growth phenotype

Supplemental table 1: Description of constructed double mutants

In supplemental table 1, all constructed mutants and their experimentally determined in vivo growth rate (in % of wild type) are given. Reasons for expected synthetically lethal phenotypes as well as potential explanations for observed in vivo growth rates are indicated

Supplementary references

54. Webb E, Claas K, Downs D: **thiBPQ encodes an ABC transporter required for transport of thiamine and thiamine pyrophosphate in Salmonella typhimurium.** *J Biol Chem* 1998, **273:**8946-8950.
55. Kustu SG, Ames GF: **The hisP protein, a known histidine transport component in Salmonella typhimurium, is also an arginine transport component.** *J Bacteriol* 1973, **116:**107-113.
56. Iyer R, Iverson TM, Accardi A, Miller C: **A biological role for prokaryotic ClC chloride channels.** *Nature* 2002, **419:**715-718.
57. Ayling PD, Mojica-a T, Klopotowski T: **Methionine transport in Salmonella typhimurium: evidence for at least one low-affinity transport system.** *J Gen Microbiol* 1979, **114:**227-246.
58. Yamada RH, Tsuji T, Nose Y: **Uptake and utilization of vitamin B6 and its phosphate esters by Escherichia coli.** *J Nutr Sci Vitaminol (Tokyo)* 1977, **23:**7-17.
59. Yamada RH, Furukawa Y: **Apparent pyridoxine transport mutants of Escherichia coli with pyridoxal kinase deficiency.** *Biochim Biophys Acta* 1980, **600:**581-584.
60. Yamada R, Furukawa Y: **Role of pyridoxal kinase in vitamin B6 uptake by Escherichia coli.** *J Nutr Sci Vitaminol (Tokyo)* 1981, **27:**177-191.
61. Meredith TC, Woodard RW: **Identification of GutQ from Escherichia coli as a D-arabinose 5-phosphate isomerase.** *J Bacteriol* 2005, **187:**6936-6942.
62. Wasserman SA, Walsh CT, Botstein D: **Two alanine racemase genes in Salmonella typhimurium that differ in structure and function.** *J Bacteriol* 1983, **153:**1439-1450.

2.4 Accumulated gene inactivation approach in *Salmonella* Typhimurium by deleting the anti-mutator genes *mutS* and *dnaQ*

(Manuscript in preparation, additional analyses are planned)

Steeb B, Schmidt A, Ahrné E, Bumann D

Summary:

Previous *in silico* results indicate that the vast majority of *Salmonella* genes is dispensable for growth. Experimental data prove the individual dispensability of metabolic genes, but experimental data about the dispensability of metabolic gene clusters is lacking.

To experimentally analyze the dispensability of metabolic gene clusters in *S.* Typhimurium, we performed an unbiased large-scale random mutagenesis approach. For this, we generated mutator lines (SL1344 *dnaQ mutS*) that possess a high mutagenesis rate. We passaged 20 separate mutator lines in two different media for 800 / 1800 generations. After complementing the mutator lines with a *dnaQ mutS* expression plasmid, we identified their genomic mutations via Illumina/Solexa deep sequencing and subsequent *in silico* analyzes. We also acquired and analyzed the proteome of all 20 lines to validate our sequence analysis findings.

Our results indicate that the mutator lines accumulated some 2700 to 3900 mutations, causing an inactivation of about 400 to 600 genes. Thus we were able to show in an unbiased large-scale mutagenesis approach that hundreds of mutations can be tolerated by *Salmonella*. We present here preliminary data that will be evaluated further in the future.

Statement of my work:

I generated all data presented in this manuscript except for the analysis of proteome data (A. Schmidt), sequencing of the *Salmonella* strains (C. Beisel, D-BSSE) and mapping *S.* Typhimurium SL1344 orthologues to *S.* Typhimurium LT2 orthologues (P. Manfredi). Plasmids with *dam*, *dnaE173* and oriT_{RP4} were obtained from M. Marinus, H. Maki and A. Böhm. I participated in writing the manuscript.

Accumulated gene inactivation approach in *Salmonella* Typhimurium by deleting the anti-mutator genes *mutS* and *dnaQ*

(Manuscript in preparation, additional analyses are planned)

Steeb B, Schmidt A, Ahrné E, Bumann D

Abstract

Previous results indicate that in *Salmonella* Typhimurium, the vast majority of metabolic genes may be not needed for replication. To experimentally analyze gene dispensability in *S*. Typhimurium, we developed an unbiased large-scale method for random mutagenesis. Deletion of the anti-mutator genes *mutS* and *dnaQ* resulted in a strain with 6.000 times elevated mutagenesis rate, which is higher than mutation rates in previous multi-cycle reductive evolution experiments. We applied this method for reductive evolution of 20 bacterial lines in two different defined media for 800 or 1800 generations. The mutator lines accumulated a total of some 66.000 mutations, with about 400 to 600 genes inactivated per line. These experimental data indicated that *Salmonella* can tolerate massive genome reductions.

Introduction

Infectious diseases are a global, rising threat [1]. Constant formation of new antimicrobial resistances leads to development of pandrug-resistant pathogens (reviewed in [2]), and new antibacterial drug targets to fight this threat are difficult to identify. Most genes in pathogens are unsuitable as drug targets, since they show only a weak phenotype during infection [3-10]. The reason for the weak phenotypes could be true dispensability of these genes, dispensability through functionally redundant enzymes / pathways or a virulence contribution below the experimental detection threshold. Genes with a minute virulence phenotype could nonetheless have an important function for virulence through a cumulative contribution to virulence (the “marginal benefit” hypothesis [11]).

In a previous study, our group analyzed the underlying reason for the predominance of non-contributing genes in the pathogen *Salmonella enterica* serovar Typhimurium (manuscript submitted). For this, we focused on metabolism as a manageable and well analyzed subsystem [12]. *In silico* and *ex vivo*

proteome analysis suggested that in a typhoid fever model only few metabolic genes have small growth rate defects below the detection threshold of some 5% [13]. Redundancy of alternative enzymes or metabolic pathways was also of minor importance during infection. In contrast to this, *in silico* analyses and *ex vivo* proteome data suggested true dispensability for the majority of metabolic enzymes, with only some 364 to 700 metabolic genes needed for replication in the typhoid fever model ([13], manuscript submitted).

Here, we wanted to experimentally validate these predictions and analyze gene dispensability in *S*.Typhimurium in an unbiased manner. For this we developed a method for random extensive genome-scale gene inactivation. To determine the feasibility of our goal to obtain highly inactivated *Salmonella* genomes, we tested the suitability of mutator genes or inactivated anti-mutator genes that might induce a higher mutation rate compared to previous mutagenesis approaches [14, 15]. Our experimental results indicated that overexpression of mutator genes was unsuitable for long term experiments due to rapid selection for mutations that diminish mutator gene function. In contrast, deletion of anti-mutator genes allowed for sustainable high mutagenesis rates and large-scale random mutagenesis, leading to rapid accumulation of extensive mutations in the *Salmonella* genome.

Results

High mutagenesis rates through expression of mutator genes selects for suppressor mutations

To increase mutation rates we first expressed a mutated form of the α-subunit of DNA polymerase III (*dnaE173*) and the DNA adenine methyltransferase *dam.* These genes are involved in chromosomal replication and the methyl directed mismatch repair system. It has been shown for both genes that an overexpression leads to increased mutation rates, including a high proportion of frameshifts [16, 17]. To use these genes for mutagenesis, we combined *dnaE173* and *dam* on a plasmid under control of the L-arabinose inducible promoter P_{BAD} (pBS12). Induction of *dam* and *dnaE173* resulted in a large range of colony sizes (see fig. 1) and strongly increased mutagenesis rates observed in the rifampicin assay (some $7x10^4$ to $7x10^5$ fold increase compared to the wildtype). However, mutagenesis rates decreased dramatically during subsequent induction cycles. Plasmid transfer experiments revealed that mutator gene functionality and / or expression were impaired after one mutagenesis cycle. Furthermore, analysis

of large colonies showed in additional cycles mutagenesis activities that were similar to separate *dam* or *dnaE173* overexpression (see Fig 2, A to I), suggesting that the other mutator gene had been impaired. This rapid failure of the mutator system prevented a practical application in multi-cycle evolution experiments.

Deletion of *mutS* and *dnaQ* leads to high mutagenesis rates

Since the inducible mutator construct proved to be unstable, we used an alternative approach for large-scale random mutagenesis. We deleted the ε-subunit of DNA polymerase III (*dnaQ*) which is responsible for proof reading during replication, and *mutS*, an enzyme essential for the methyl directed mismatch repair system. Deletion of these anti-mutator genes led to a similarly heterogeneous colony size distribution as observed for *dnaE173* and *dam* overexpression (see fig.4A). SL1344 *dnaQ mutS* demonstrated a very high mutation rate which was observed in the rifampicin assay (see fig. 3). 20 separate mutator lines were passaged for some 800 to 1800 generations in minimal or rich media. The lines were then complemented with an inducible *dnaQ mutS* expression plasmid (pBS33), which restored growth and low mutation rates (see fig. 3, fig. 4B). We continued to work with these stabilized lines.

Mutation spectrum and mutation rates of *mutS dnaQ* null mutants

We identified genome sequences of all 20 stabilized mutator lines via Solexa deep sequencing. We detected in total some 66.000 mutations, with a ratio of single nucleotide polymorphisms (SNPs) to insertions / deletions (InDels) of two to one. We mapped these 66.000 mutations on the chromosome of *S*. Typhimurium SL1344 to determine mutagenesis coverage. The average distance between mutations was 74 bp ± 117 bp, the maximal detected distance between two mutations was 6246 bp.

SNPs did not show a preference for hotspot sequences and only a slight preference for GC over AT mutations. Transitions made up 95% of all SNP mutations (see tab. 1). In comparison to this, InDels showed a strong tendency for AT over GC mutations (see tab. 1) as well as for sequence hotspots, with insertions being more sensitive to homopolynucleotide runs compared to deletions (Insertions: 91 ± 4%

in runs of at least four identical nucleotides; deletions: 68 ± 15% in runs of at least four identical nucleotides). The majority (>98%) of InDels affected one to two bases. Based on the estimated amount of generations during the experiment, the mutagenesis rate was in the range of 2.2 to 3 mutations per division (see tab. 2).

Gene inactivation by random mutagenesis

We were interested in which genes can be inactivated in the *Salmonella* genome. We considered only InDels and nonsense-mutations within the central 60% of a gene as inactivating mutations to minimize the risk of false positives (see Materials and Methods for a more extensive explanation).
The average number of inactivated genes per mutator line was some 400 to 600 genes. In total, 2976 genes were inactivated at least once in one of the 20 lines, which represented two thirds of the total number of genes in the genome of *S.* Typhimurium SL1344 (see fig. 5).

Comparison to other data sets

We compared genes inactivated in our mutagenesis approach with data for a defined single gene knockout library of *E. coli* (Keio collection [18]). In 12 out of 1732 cases, our experimental data differed from the Keio collection. These conflicts could be explained by rich nutritional supplementation in our approach (*purB*). In some cases, viable mutants of putative essential genes were also reported in other literature sources (*yraL* [19], *ftsN* [20], *minD* [21]*, lpxB* [22], *rfaK* [23], *entD* [24], *bscB* [25] and *yrfF* [26]). Thus, from 12 conflicts, only 3 could not be explained (*yqgD*, *yhhQ*, *ydiL*), which was due to lacking information about the corresponding enzymes.

Massive mutagenesis leads to distinct phenotypes

All mutator strains had mild to severe growth rate defects. On average, rich medium mutator lines grew at 50% of the wildtype rate, whereas minimal medium lines had some 70% of the wildtype growth rate (see fig. 6).

Preliminary proteome analysis of the mutator lines indicated drastic differences in the protein content of the mutator lines. On average, 32 ± 13 proteins were more than tenfold expressed compared to the wildtype, 215 ± 43 proteins were less than 0.1 fold expressed compared to the wildtype (see fig. 7). Preliminary analysis also indicated a higher concentration of GroEL/GroES and DnaK, especially in heavily mutated and fast growing rich medium lines.

Metabolic analysis of genes inactivated in the mutator lines revealed that diverse catabolic pathways were inactivated in the various lines (glycolysis, β-oxidation, parts of mixed acid fermentation, glyoxylate cycle, degradation of diverse carbohydrates). Specifically in lines grown in rich media, multiple synthesis reactions of amino acids and nucleosides were inactivated (see fig. 8).

Discussion

To analyze gene dispensability in *Salmonella* in an unbiased manner, we tested the suitability of different approaches for random mutagenesis. Other groups were able to show that overexpression of mutator genes or deletion of anti-mutator genes can be used for random mutagenesis [14, 15, 27]. We were interested in higher mutagenesis rates than reported in these experiments to allow for a more rapid reductive evolution process. Our results indicated that overexpression of genes with high mutagenesis rate are unsuitable for long-term experiments. In contrast to this, the deletion of the anti-mutator genes *dnaQ* and *mutS* was successfully used for a reductive evolution approach, enhancing the mutagenesis rate sustainably to a higher level. Passaging of the *dnaQ mutS* mutator strain led to massive accumulation of mutations in the *Salmonella* genome.

In a first approach, we overexpressed the mutator genes *dnaE173* and *dam*. This resulted in a high mutagenesis rate in the first cycle, but mutations in the mutator plasmid diminished mutation rates in subsequent cycles. A decreased mutagenesis rate likely caused fitness advantages, which resulted in rapid displacement of clones that retained fully functional mutator genotypes. The overexpression of mutator genes was thus not appropriate for large-scale mutagenesis in multiple cycles. Appearance of suppressor mutations during overexpression of strong mutator genes was previously observed by other groups. Maisnier-Patin *et al.* demonstrated that overexpression of *dinB* can be used for random mutagenesis, resulting in a mutagenesis rate of up to 230 fold compared to the wildtype. In line with our results, three from eight mutator lines with highest *dinB* expression showed in the course of their experiment altered mutagenesis activity due to suppressor mutations that occurred in *dinB* [27].

Our aim was to develop an approach for large-scale random mutagenesis for fast inactivation of numerous genes in *Salmonella*. This multi-cycle mutagenesis was not achievable with overexpression of *dnaE173* and *dam* due to rapid inactivation. We therefore chose to delete the antimutator genes *dnaQ* and *mutS*, which participate in the same cellular mechanisms of replication and methyl directed mismatch repair system as *dnaE173* and *dam*. After passaging of the mutator lines for 800 to 1800 generations, SL1344 *dnaQ mutS* lines were complemented with a plasmid containing *mutS* and *dnaQ* under control of the arabinose inducible promoter P_{BAD} to reduce mutagenesis extent and to stabilize their genome.

Subsequent whole genome deep sequencing revealed the extent of mutagenesis. Deletion of *mutS* and *dnaQ* led to a 6.000 fold higher mutagenesis rate compared to the wildtype, and 20 to 26 fold higher mutagenesis rates than obtained by *mutS* deletion or *dinB* overexpression used in previous multi-cycle reductive evolution experiments [15, 27, 28]. On the other hand, the very high mutagenesis rate of the SL1344 *dnaQ mutS* strain led to frequent cell death through mutation of essential genes, which reduced the net growth rate of the strain, rendering it most likely avirulent in host tissues.

As expected, lines that were passaged for more generations due to faster growth in rich medium accumulated more mutations than lines that were passaged for about half the number of generations in minimal medium. Nonetheless, minimal medium lines showed a higher mutagenesis rate per generation. A possible explanation could be that the number of divisions was underestimated in cells grown in minimal media. Since 114 genes are required for prototrophic growth (based on the Keio collection [18]), mutagenesis in minimal media could lead to a higher amount of cell death and cell lysis compared to

mutagenesis in rich media. Cryptic growth of surviving bacteria would then result in a higher absolute number of cell divisions with the same final growth yield [29], leading to an overestimation of the overall mutagenesis rate per division.

All in all, our results demonstrated that deletion of the anti-mutator genes *dnaQ* and *mutS* can be used for unbiased large-scale mutagenesis. We were able to show that *Salmonella* has the capacity to tolerate a massive amount of mutations *in vitro* and that this mutagenesis leads to drastic proteome changes. Future efforts will be focused on a detailed analysis of genotypic and phenotypic characterization of all mutator lines.

Materials and Methods

Growth media. Rich and minimal medium was based on modified M9 minimal medium.

Minimal medium contained instead of glucose only glycerol [0.4 M] as carbon source and 0.0002% L-arabinose for induction of pBS33 based expression of *dnaQ* and *mutS*.

Rich medium contained instead of glucose only glycerol [0.4 M] as carbon source and 0.0002% L-arabinose for induction of pBS33 based expression of *dnaQ* and *mutS*. Rich medium also contained various nucleosides (Thymidine [300 µM], uridine [300 µM], guanosine [300 µM], cytidine [300 µM], adenosine [300 µM] and 2-deoxy-D-ribose [300 µM]). Amino acids were supplemented via RPMI 1640 Amino acids solution (50x, Sigma) [2.5x final concentration] and separate addition of L-alanine [250 µM] and L-glutamine [250 µM]. Vitamins were also added (RPMI 1640 Vitamins Solution (100x, Sigma) [1x final concentration]).

If needed, we used the following antibiotica in their indicated concentration: Ampicillin [100 µg/ml], streptomycin [90 µg/ml], chloramphenicol [20 µg/ml], kanamycin [50 µg/ml] and rifampicin [100 µg/ml].

Passaging of bacteria. We used strain *Salmonella enterica* serovar Typhimurium SL1344 *hisG xyl* [30] as parental wildtype strain. Bacteria were passaged mainly in liquid media. Minimal medium lines were diluted 1:1000 every 2 days, mutator lines in rich medium were diluted 1:100 each day in the morning and 1:1000 in the evening. In absence of growth after incubation, higher amounts of previous cultures were used for fresh inoculation. Every four weeks, cultures were plated out to select for growth

capability on solid media and used for the next liquid media passage. Bacteria were stored at -80°C at regular intervals.

Calculation of divisions accumulated during mutator line passaging. Colonies on plates were estimated to be formed by 23.25 divisions (10^7 CFU). Dilution in liquid media was estimated to lead to 6.6 divisions (1:100) respectively 10 divisions (1:1000).

Generation of mutants. Mutants were constructed by lambda red-recombinase mediated allelic replacement [31].

Construction of plasmids for mutagenesis, complementation and conjugation. The genes *dnaE173* and *dam* were obtained from plasmids pNF5-*dnaE173* [32] and pMQ400 [33] and cloned into plasmid pMW279 under control of its P_{BAD} promoter (pBS12). The $oriT_{RP4}$ region of plasmid pLD54 [34] was introduced into plasmid pMW279 to allow for its conjugational transfer (pBS20). *dnaQ* and *mutS* from *Salmonella* Typhimurium was cloned into pBS20 under the control of its P_{BAD} promoter (pBS33).

Conjugation. Conjugation was done based on the protocol of Herrero *et al.* [35] and Demarre *et al.* [36]. In brief, the plasmid containing donor strain *E. coli* BW20767 [34] and the recipient strain were picked from plate and mixed in a ratio of 1:10 in PBS or LB. 200 µl of this solution was used to imbue a filter disc (Millipore) on an LB plate without antibiotics, which was then incubated at 37°C for 8 hours. Then the filter disc was immersed in 10 mM MgSO4 solution, which was diluted and plated out on LB plates containing streptomycin and ampicillin to select for ex-conjugants. Successful conjugation was verified by colony PCR, testing for the presence of the plasmid in combination with the presence of the *Salmonella* specific gene *srfJ*.

Rifampicin assay. Overnight cultures of bacteria were plated out on LB / rich media plates to determine the CFU per ml and on LB / rich media plates with rifampicin to determine the amount of Rif^r per 10^8 CFU. Plates were incubated for 48 h before analysis.

Growth rate determination. Erlenmeyer beakers with bacteria were incubated in a water bath at 37°C to minimize temperature fluctuations. The growth rates were determined in duplicates and displayed as average value of both measurements.

Proteome analysis. 10 ml of overnight culture were pelleted and washed once with PBS. The cell pellet was lysed with 8M urea and RapiGest (0.1%) in combination with ultra sonification. Protein concentration was determined via a BCA assay (thermo scientific pierce BCA protein assay kit) according to the manufacturer's description. 100 µg protein was reduced and alkylated via addition of tris(2-carboxyethyl)phosphine (TCEP), iodoacetamide and N-acetyl-cysteine. Proteins were then digested with the proteases Lys-C and trypsin. After digestion, a mix of heavy labeled reference AQUA peptides was added to wildtype samples for peptide quantification. Peptides were bound to C18 columns (Macro SpinColumns, Harvard Apparatus) and washed with acetonitrile and trifluoroacetic acid (TFA) containing buffers. The eluated peptide mixture was dried and resuspended in LC-MS/MS buffer in a final concentration of 0.5 µg/µl. Samples were analysed by LC-MS/MS using an LTQ-Orbitrap-Velos instrument (Thermo-Fischer Scientific).

Whole genome deep sequencing. Genomic DNA was prepared out of the bacterial amount corresponding to 12 OD of an overnight culture, using the Genelute Bacterial genomic DNA Kit (Sigma) according to manufacturer's description. Samples were then further processed at the ETH Zürich Department of Biosystems Science and Engineering in Basel (D-BSSE). In brief, 50 ng of genomic DNA was treated with the Nextera™ DNA Sample Prep Kit (Epicentre) to prepare the samples for sequencing. Samples were sequenced in 50 cycles with the Solexa method (Illumina).

In silico analysis. Solexa sequence output was prepared for analysis by in-house Perl algorithms and analyzed via VAAL (v1.2) [37]. The VAAL output was then processed by diverse in-house Perl algorithms to identify mutated genes. Perl algorithms were based upon code presented in www.szabgab.com, www.perldoc.perl.org, www.tizag.com, www.perlmonks.org and [38]. Genome annotation information and genome sequences were downloaded from the homepage http://brcdownloads.vbi.vt.edu/patric2/genomes/.

Gene inactivity evaluation. Only InDels with a size unequal of 3 or 6 nucleotides and nonsense mutations were evaluated for a possible gene inactivation. Any of these mutations within the arbitrary threshold of 20 to 80% of the open reading frame (ORF) were designated to cause a gene inactivation. A mutation within the first 20% of the ORF was also designated as a cause for gene inactivation, if no downstream methionine was encoded within the residual part of the 20% that could serve as alternative start codon.

Ortholog identification. To enable comparison of the mutagenesis results of *Salmonella* Typhimurium SL1344 with the highly annotated *Salmonella* Typhimurium LT2 genome, mapping of SL1344 orthologs with LT2 genes was done with OrthoMCL (v1.4) [39]. Shared synteny between both *Salmonella* serovars was used to identify the correct orthologous pairs, if multiple combinations were possible.

Acknowledgement

We thank P. Manfredi for mapping *S*.Typhimurium SL1344 orthologues to *S*.Typhimurium LT2 orthologues and C. Beisel (D-BSSE) for executing the deep sequencing. We thank M. Marinus and H. Maki for supplying plasmids with *dam* and *dnaE173.* We thank A. Böhm for supply of the strains BW20767 and BW21038 pLD54, used for extraction of the oriT$_{RP4}$ and for conjugation.

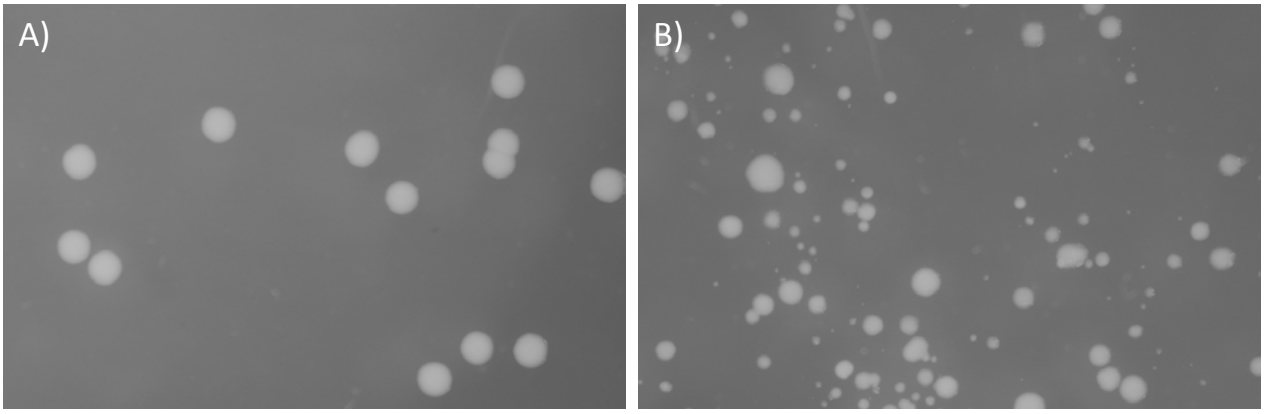
References

1. Norrby SR, Nord CE, Finch R: **Lack of development of new antimicrobial drugs: a potential serious threat to public health**. *Lancet Infect Dis* 2005, **5**(2):115-119.
2. Falagas ME, Bliziotis IA: **Pandrug-resistant Gram-negative bacteria: the dawn of the post-antibiotic era?** *Int J Antimicrob Agents* 2007, **29**(6):630-636.
3. Chaudhuri RR, Peters SE, Pleasance SJ, Northen H, Willers C, Paterson GK, Cone DB, Allen AG, Owen PJ, Shalom G *et al*: **Comprehensive identification of Salmonella enterica serovar typhimurium genes required for infection of BALB/c mice**. *PLoS Pathog* 2009, **5**(7):e1000529.
4. Santiviago CA, Reynolds MM, Porwollik S, Choi SH, Long F, Andrews-Polymenis HL, McClelland M: **Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice**. *PLoS Pathog* 2009, **5**(7):e1000477.
5. Kavermann H, Burns BP, Angermuller K, Odenbreit S, Fischer W, Melchers K, Haas R: **Identification and characterization of Helicobacter pylori genes essential for gastric colonization**. *J Exp Med* 2003, **197**(7):813-822.
6. Maroncle N, Balestrino D, Rich C, Forestier C: **Identification of Klebsiella pneumoniae genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis**. *Infect Immun* 2002, **70**(8):4729-4734.
7. Edelstein PH, Edelstein MA, Higa F, Falkow S: **Discovery of virulence genes of Legionella pneumophila by using signature tagged mutagenesis in a guinea pig pneumonia model**. *Proc Natl Acad Sci U S A* 1999, **96**(14):8190-8195.
8. Sun YH, Bakshi S, Chalmers R, Tang CM: **Functional genomics of Neisseria meningitidis pathogenesis**. *Nat Med* 2000, **6**(11):1269-1273.
9. Chiang SL, Mekalanos JJ: **Use of signature-tagged transposon mutagenesis to identify Vibrio cholerae genes critical for colonization**. *Mol Microbiol* 1998, **27**(4):797-805.
10. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C: **Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis**. *Mol Microbiol* 1999, **34**(2):257-267.

11. Thatcher JW, Shaw JM, Dickinson WJ: **Marginal fitness contributions of nonessential genes in yeast**. *Proc Natl Acad Sci U S A* 1998, **95**(1):253-257.
12. Bumann D: **System-level analysis of Salmonella metabolism during infection**. *Curr Opin Microbiol* 2009, **12**(5):559-567.
13. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, Mann M, Bumann D: **Robust Salmonella metabolism limits possibilities for new antimicrobials**. *Nature* 2006, **440**(7082):303-307.
14. Nilsson AI, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JC, Andersson DI: **Bacterial genome size reduction by experimental evolution**. *Proc Natl Acad Sci U S A* 2005, **102**(34):12112-12116.
15. Funchain P, Yeung A, Stewart JL, Lin R, Slupska MM, Miller JH: **The consequences of growth of a mutator strain of Escherichia coli as measured by loss of function among multiple gene targets and loss of fitness**. *Genetics* 2000, **154**(3):959-970.
16. Yang H, Wolff E, Kim M, Diep A, Miller JH: **Identification of mutator genes and mutational pathways in Escherichia coli using a multicopy cloning approach**. *Mol Microbiol* 2004, **53**(1):283-295.
17. Mo JY, Maki H, Sekiguchi M: **Mutational specificity of the dnaE173 mutator associated with a defect in the catalytic subunit of DNA polymerase III of Escherichia coli**. *J Mol Biol* 1991, **222**(4):925-936.
18. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: **Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection**. *Mol Syst Biol* 2006, **2**:2006 0008.
19. Dassain M, Leroy A, Colosetti L, Carole S, Bouche JP: **A new essential gene of the 'minimal genome' affecting cell division**. *Biochimie* 1999, **81**(8-9):889-895.
20. Dai K, Xu Y, Lutkenhaus J: **Cloning and characterization of ftsN, an essential cell division gene in Escherichia coli isolated as a multicopy suppressor of ftsA12(Ts)**. *J Bacteriol* 1993, **175**(12):3790-3797.
21. Parti RP, Biswas D, Wang M, Liao M, Dillon JA: **A minD mutant of enterohemorrhagic E. coli O157:H7 has reduced adherence to human epithelial cells**. *Microb Pathog* 2011, **51**(5):378-383.
22. Bulawa CE, Raetz CR: **The biosynthesis of gram-negative endotoxin. Identification and function of UDP-2,3-diacylglucosamine in Escherichia coli**. *J Biol Chem* 1984, **259**(8):4846-4851.
23. MacLachlan PR, Kadam SK, Sanderson KE: **Cloning, characterization, and DNA sequence of the rfaLK region for lipopolysaccharide synthesis in Salmonella typhimurium LT2**. *J Bacteriol* 1991, **173**(22):7151-7163.
24. Fleming TP, Nahlik MS, Neilands JB, McIntosh MA: **Physical and genetic characterization of cloned enterobactin genomic sequences from Escherichia coli K-12**. *Gene* 1985, **34**(1):47-54.
25. Ledeboer NA, Jones BD: **Exopolysaccharide sugars contribute to biofilm formation by Salmonella enterica serovar typhimurium on HEp-2 cells and chicken intestinal epithelium**. *J Bacteriol* 2005, **187**(9):3214-3226.
26. Costa CS, Pettinari MJ, Mendez BS, Anton DN: **Null mutations in the essential gene yrfF (mucM) are not lethal in rcsB, yojN or rcsC strains of Salmonella enterica serovar Typhimurium**. *FEMS Microbiol Lett* 2003, **222**(1):25-32.
27. Maisnier-Patin S, Roth JR, Fredriksson A, Nystrom T, Berg OG, Andersson DI: **Genomic buffering mitigates the effects of deleterious mutations in bacteria**. *Nat Genet* 2005, **37**(12):1376-1379.
28. Wielgoss S, Barrick JE, Tenaillon O, Cruveiller S, Chane-Woon-Ming B, Medigue C, Lenski RE, Schneider D: **Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With Escherichia coli**. *G3 (Bethesda)* 2011, **1**(3):183-186.
29. Druilhet RE, Sobek JM: **Starvation survival of Salmonella enteritidis**. *J Bacteriol* 1976, **125**(1):119-124.

30. Hoiseth SK, Stocker BA: **Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines**. *Nature* 1981, **291**(5812):238-239.
31. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products**. *Proc Natl Acad Sci U S A* 2000, **97**(12):6640-6645.
32. Yanagihara F, Yoshida S, Sugaya Y, Maki H: **The dnaE173 mutator mutation confers on the alpha subunit of Escherichia coli DNA polymerase III a capacity for highly processive DNA synthesis and stable binding to primer/template DNA**. *Genes Genet Syst* 2007, **82**(4):273-280.
33. Calmann MA, Marinus MG: **Regulated expression of the Escherichia coli dam gene**. *J Bacteriol* 2003, **185**(16):5012-5014.
34. Metcalf WW, Jiang W, Daniels LL, Kim SK, Haldimann A, Wanner BL: **Conditionally replicative and conjugative plasmids carrying lacZ alpha for cloning, mutagenesis, and allele replacement in bacteria**. *Plasmid* 1996, **35**(1):1-13.
35. Herrero M, de Lorenzo V, Timmis KN: **Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria**. *J Bacteriol* 1990, **172**(11):6557-6567.
36. Demarre G, Guerout AM, Matsumoto-Mashimo C, Rowe-Magnus DA, Marliere P, Mazel D: **A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries and their cognate Escherichia coli host strains**. *Res Microbiol* 2005, **156**(2):245-255.
37. Nusbaum C, Ohsumi TK, Gomez J, Aquadro J, Victor TC, Warren RM, Hung DT, Birren BW, Lander ES, Jaffe DB: **Sensitive, specific polymorphism discovery in bacteria using massively parallel sequencing**. *Nat Methods* 2009, **6**(1):67-69.
38. Tisdall JD: **Beginning Perl for bioinformatics**, 1st edn. Beijing ; Sebastopol, CA: O'Reilly; 2001.
39. Li L, Stoeckert CJ, Jr., Roos DS: **OrthoMCL: identification of ortholog groups for eukaryotic genomes**. *Genome Res* 2003, **13**(9):2178-2189.
40. Paley SM, Karp PD: **The Pathway Tools cellular overview diagram and Omics Viewer**. *Nucleic Acids Res* 2006, **34**(13):3771-3778.

Figures



SL1344 SL1344 pBS12 (*dnaE173 dam*)

Fig. 1: Colony morphology of wildtype SL1344 **(A)** and SL1344 pBS12 (*dnaE173 dam*) after induction **(B)**. Whereas the colony size of the wildtype is homogeneous, a single overexpression cycle of the mutator genes *dnaE173* and *dam* leads to a variation in colony size and colony morphology.

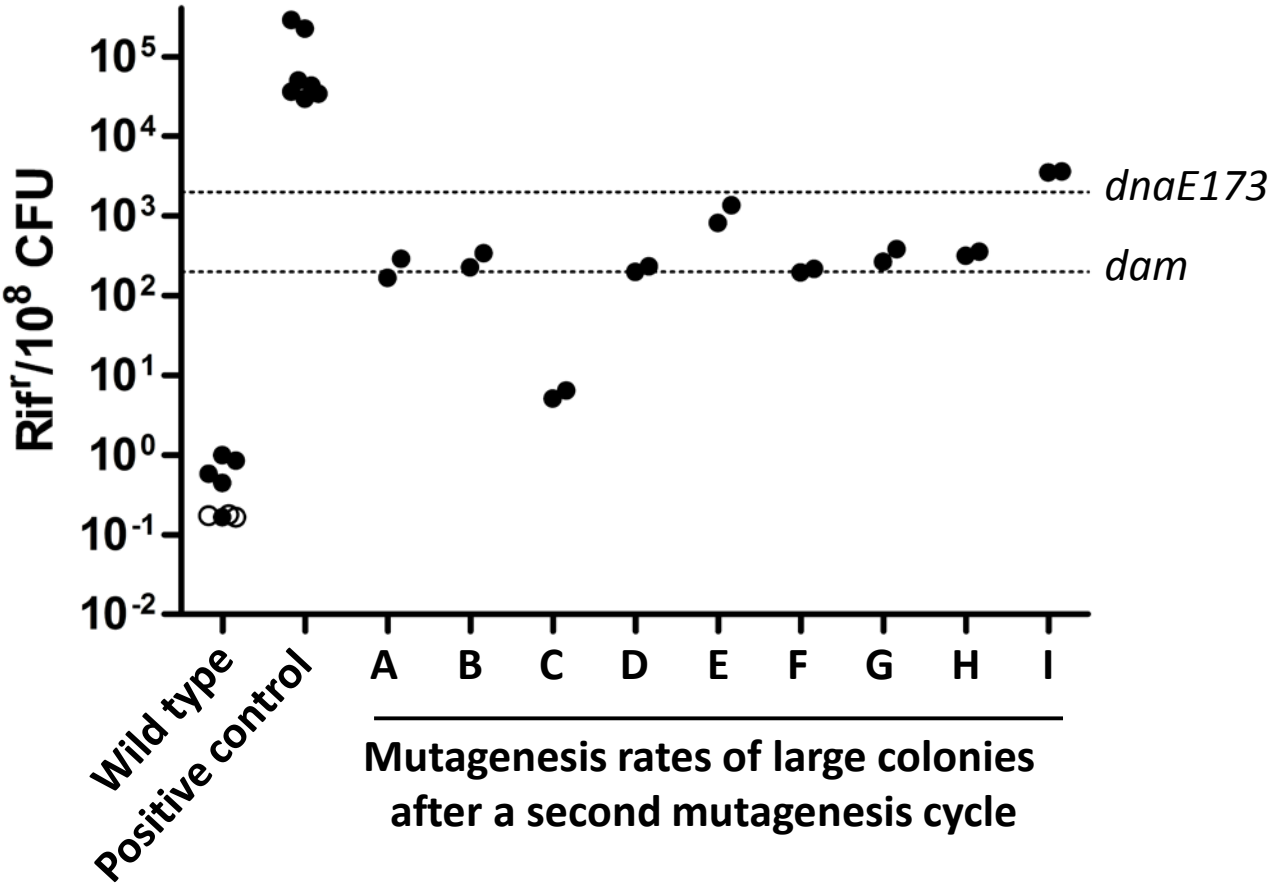


Fig. 2: Rifampicin resistance (Rif^r) after one cycle of *dnaE173* and *dam* mediated mutagenesis (positive control) or of large colonies subject to a second mutagenesis cycle (A to I).
Positive control and wildtype control data were pooled from two subsequent experiments. For the wildtype, three data points were below the detection threshold of some 0.15 Rif^r /10^8 CFU, indicated as open circles. Values for average amount of Rif^r per 10^8 CFU for *dam* (some 200 Rif^r per 10^8 CFU) and *dnaE173* (some 2000 Rif^r per 10^8 CFU) were obtained in independent experiments.

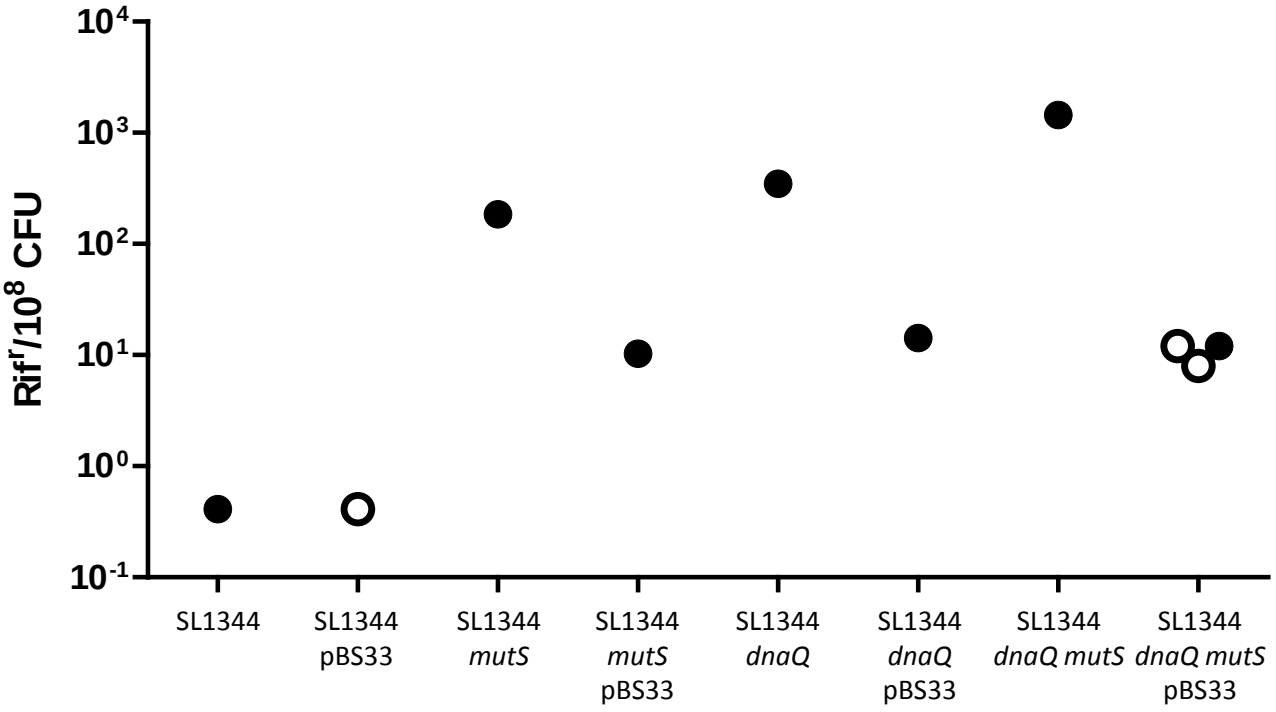


Fig.3: Rifampicin resistance in SL1344 (wildtype) and different mutator lines with and without *dnaQ mutS* complementation.
For the complemented wildtype and mutator strain SL1344 *dnaQ mutS*, data points were below the detection threshold (some 0.4 Rif^r /10^8 CFU for SL1344 pBS33, some 10 Rif^r /10^8 CFU for SL1344 *dnaQ mutS* pBS33), indicated as open circles in fig. 3.

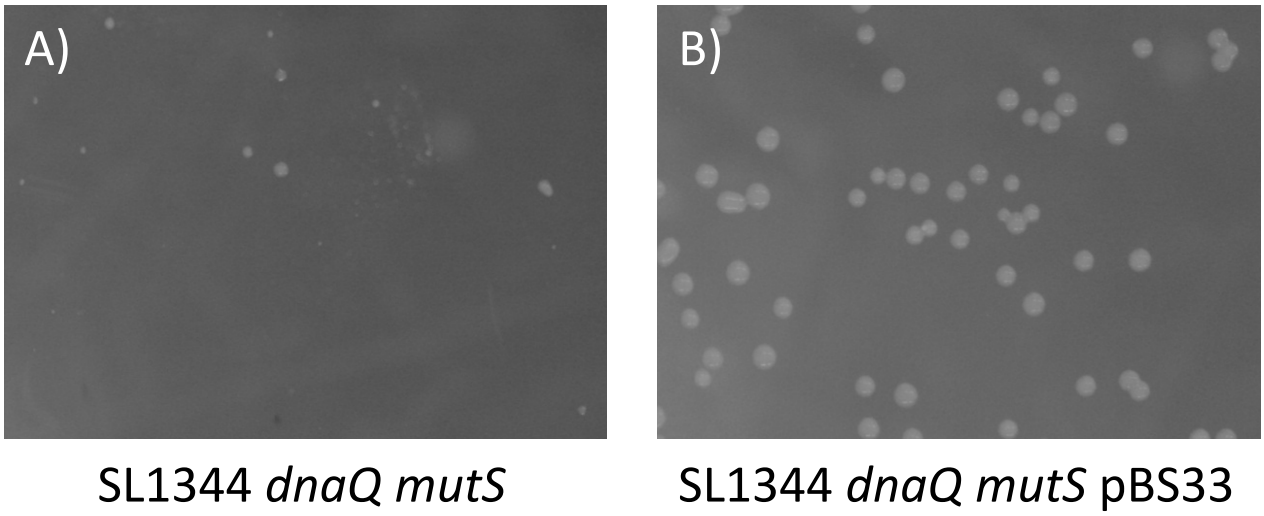


Fig. 4: Colony morphology of SL1344 *dnaQ mutS* **(A)** and SL1344 *dnaQ mutS* pBS33 (*dnaQ mutS*) after induction **(B)**.
Whereas the mutator line shows inhomogeneous colony morphology, complementation and induction of *dnaQ* and *mutS* led to a higher growth yield and more homogeneous colony size and morphology. The photo was taken after overnight incubation.

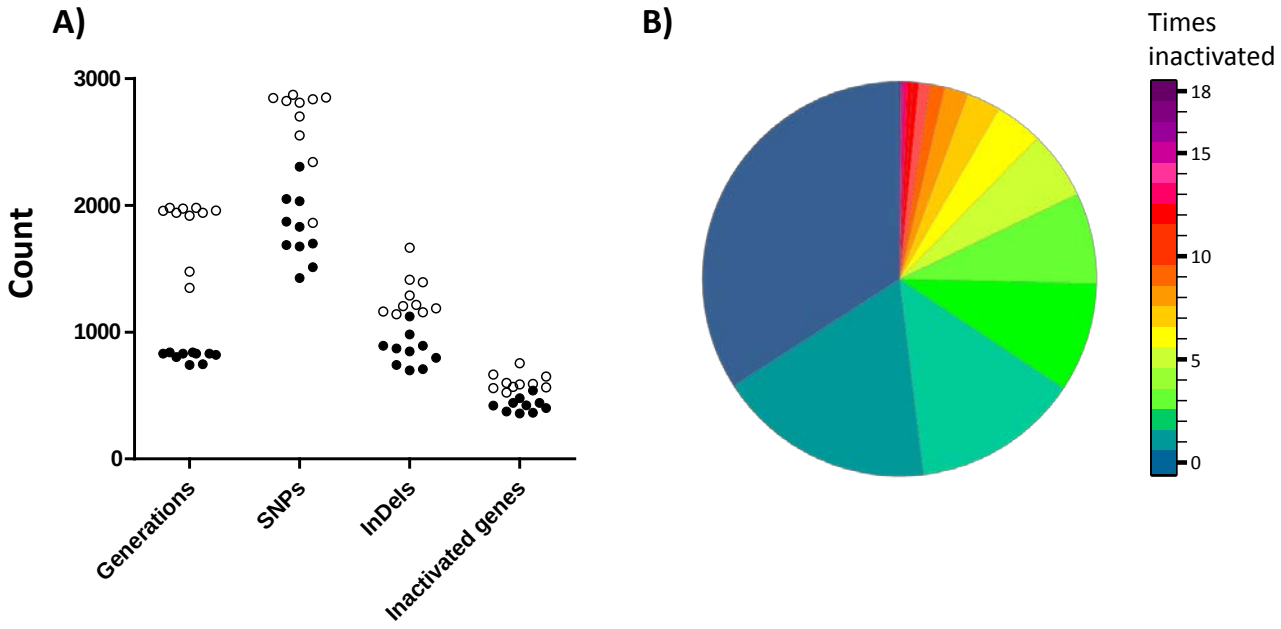


Fig. 5: Number of mutations and inactivated genes in mutator lines grown in rich and minimal media. **(A)** Overview about mutagenesis in rich media (open circles) and minimal media (filled circles). **(B)** About two thirds of all genes in the SL1344 genome were inactivated at least once in one of the 20 mutator lines. The legend indicates in how many lines a gene was inactivated.

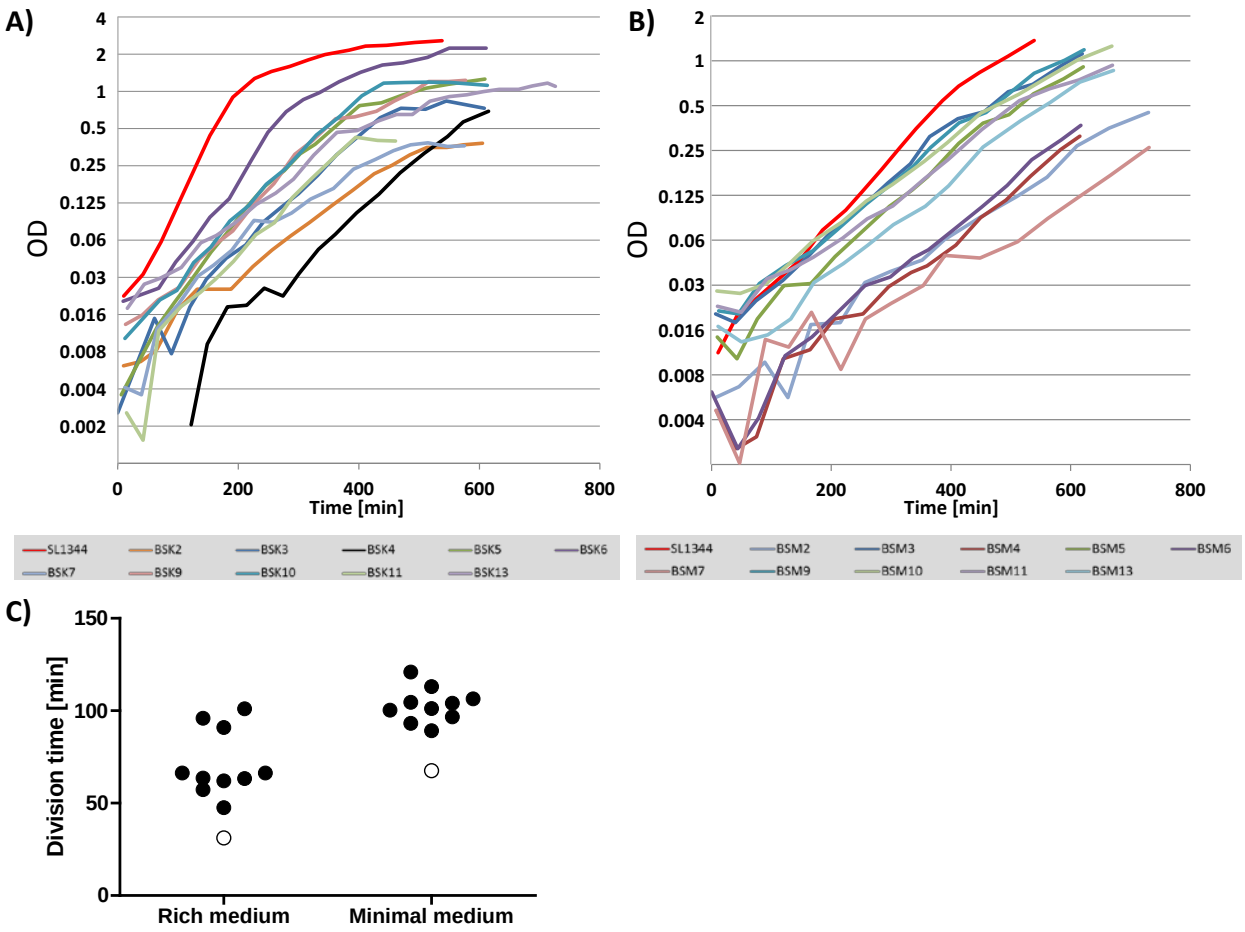


Fig. 6: Growth of SL1344 and the complemented mutator lines in rich and minimal media.
(A) Growth curves of complemented mutator lines passaged in rich medium compared to the wildtype (red line). **(B)** Growth curve of complemented mutator lines passaged in minimal medium compared to the wildtype (red line). **(C)** Division times of mutator lines passaged in rich media and in minimal media (filled circles) compared to the wildtype (open circles).
Shown data is based on the average value of two biological replicates. Exponential growth for the determination of division times was estimated in the OD range 0.0625-0.275 for rich medium lines and 0.0625-0.5 for minimal medium lines.

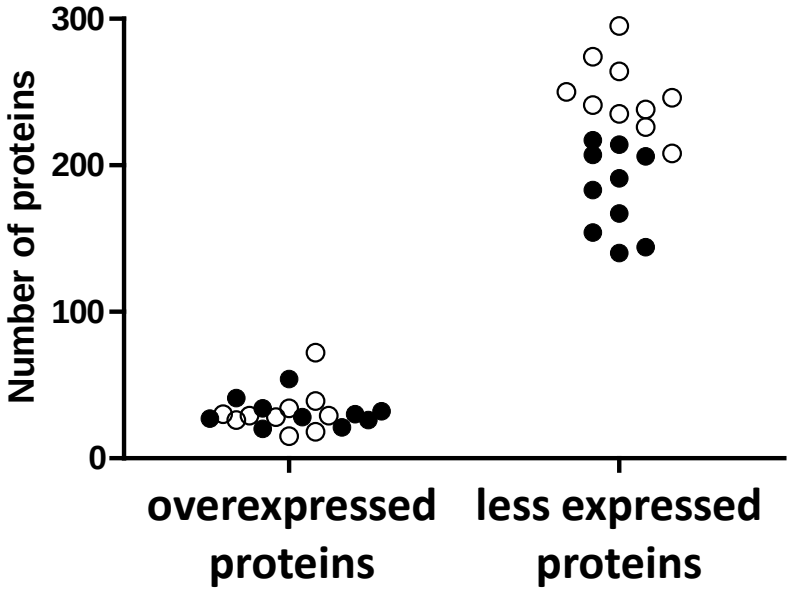


Fig. 7: Proteome changes of mutator lines passaged in rich media (open circles) and minimal media (filled circles) compared to the wildtype. Indicated are the numbers of proteins with > 10 fold overexpression or < 0.1 fold lower expression compared to wildtype grown in the same media.

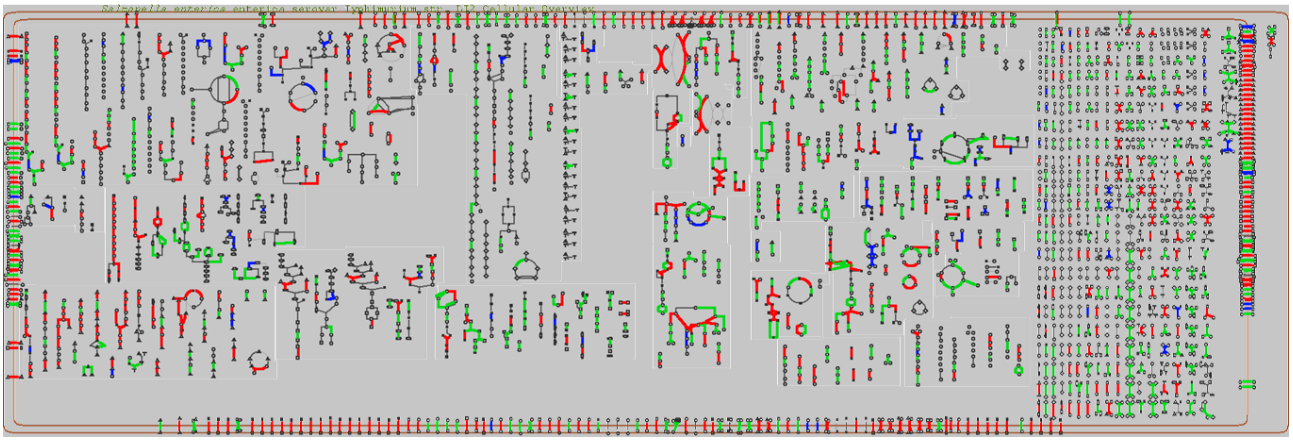


Fig. 8: Map of *Salmonella* metabolism.
In this chart, we displayed metabolic reactions associated with genes inactivated in mutator lines grown in rich media (**red**), in minimal media (**blue**) and genes inactivated in both conditions (**green**). The chart was generated with the pathway tools software package [40].

Tables

Tab. 1: Mutation spectrum of SL1344 *dnaQ mutS*.
The mutation spectrum was determined by genome sequencing of the passaged mutator lines and subsequent *in silico* analysis with VAAL [37]. Data was pooled from all 20 lines grown in minimal and rich media.

	Insertions	Deletions	SNPs	% Transition	% Transversion
AT	6272	8343	20465	95.9	4.1
GC	1787	4890	24107	94.6	5.4
% AT	77.8	63	45.9	-	-

Tab. 2: Mutation results per mutator line
The mutation results per line were determined by genome sequencing of the passaged mutator lines and subsequent *in silico* analysis with VAAL [37]. Data was pooled from all 10 lines grown in the same medium.

	SNPs	InDels	Divisions	SNPs / Division	InDels / Division
Rich medium lines	2649 ± 324	1282 ± 165	1847 ± 232	1.46 ± 0.29	0.71 ± 0.13
Minimal medium lines	1809 ± 266	855 ± 131	811 ± 37	2 ± 0.37	1 ± 0.19

3. Discussion

Salmonella in vivo metabolism is of crucial importance for its virulence, but a comprehensive and quantitative understanding of metabolism during infection has been lacking. To address this issue we used an integrated experimental / *in silico* approach. First, we generated an *in silico* reconstruction of *Salmonella* metabolism (see 2.1). To adapt this reconstruction to *in vivo* conditions, we used *Salmonella in vivo* mutant phenotypes that were informative for host nutrient supply (see 2.2). The resulting *in vivo* model of *Salmonella* metabolism allowed predicting *in vivo* nutrient fluxes and nutrient utilization mechanisms. These results indicated a qualitatively nutrient rich, but quantitatively nutrient poor environment, supporting only slow nutrient-limited *Salmonella* growth. Furthermore, the *in vivo* model allowed analyzing the contribution of genes with undetectable phenotype to virulence (see 2.3). Our results suggest that most genes were truly dispensable for replication *in vivo* because of availability of diverse host nutrients and extensive metabolic redundancy. For experimental validation of massive dispensability, methods for large scale gene inactivation approaches were developed. Preliminary results indicate the possibility to inactivate a major portion of the *Salmonella* genome (see 2.4).

3.1 Reconstruction of *Salmonella* metabolism

To enable a quantitative understanding of metabolism through *in silico* analysis, we first generated an *in silico* reconstruction of *Salmonella* metabolism. Metabolic reconstructions (MRs) are a mathematical representation of metabolism, containing all known metabolic reactions with associated enzymes and metabolites in their respective stoichiometry. MR allow *in silico* analyses of metabolism in combination with flux balance analysis (FBA) (reviewed in [85]). Several pathogen MRs have been generated and analyzed with FBA [86-89], but a reconstruction of *S.* Typhimurium was not available at the beginning of this thesis.

To generate a MR of *Salmonella* metabolism, we used an already published MR of *E. coli* (iAF1260) as a "backbone" for our reconstruction [81]. Due to their relatively recent separation some 120 to 160 million years ago, about 70% of *Salmonella* genes have similar *E. coli* orthologs [90, 91]. We retained 1075 ortholog enzymes from iAF1260 in our reconstruction and removed 185 *E. coli* specific enzymes.

Removal of pathways was in agreement with physiological data and included utilization of D-lactose, D-allose, D-galacturonic acid and sucrose ([77, 81], see Biolog's Phenotyping MicroArray™ technology website (http://www.biolog.com)). Based on literature mining, we included 145 additional *Salmonella* enzymes into the reconstruction, including salmochelin synthesis [92-96] and catabolism of metabolites such as myo-inositol or 4-hydroxyphenylacetate [97-99]. The resulting metabolic reconstruction BRecon contained 1220 genes.

MRs like BRecon can be further improved when new experimental data becomes available. Moreover, the content of reconstructions is influenced by somewhat subjective choices in literature research and interpretation [1]. To improve our reconstruction, we therefore co-organized a jamboree to merge BRecon with the independently generated *Salmonella* MR AJrecon (a variant is published in [100]) to a consensus reconstruction of *Salmonella* metabolism (STMV1.0) [1]. STMv1.0 includes 1270 genes, 2201 metabolic reactions, three separate compartments (cytosol, periplasm and extracellular), and is one of the largest and most recent MR for prokaryotes up to now [86, 101-103]. During the generation of STMv1.0, a third and smaller *Salmonella* MR was published (iMA945) [104]. This reconstruction was not included into STMv1.0, since most peculiarities of iMA945 lacked bibliomic support.

In combination with FBA, STMv1.0 allowed a comprehensive analysis of *Salmonella in vitro* metabolism. FBA calculations predicted *Salmonella* growth on diverse compounds with an overall accuracy of 88% [1], which is comparable with other recent MRs [81, 86, 105]. MR like STMv1.0 were used for several other metabolic *in silico* analyses, such as simulation of antibiotic effects [100], gene essentiality and phenotype prediction for single or multiple gene deletions [83, 106, 107], metabolic network activity prediction [108] and nutrient exchange flux determination [109]. Analysis of MRs can also be used to systematically reveal knowledge gaps between experimental data and *in silico* predictions, or highlight metabolic reactions that lack an enzyme annotation. Our MR of *Salmonella* metabolism revealed that for about 9% of all included metabolic reactions the catalyzing enzyme is not known. A high number of these non-annotated reactions was associated with metabolite transport across the inner membrane. Transporter annotation is in general lacking, since transporter identification can be hampered through redundancy as well as broad specificity of transporters [110-113]. As seen for the prediction of double mutants (see 2.3), knowledge gaps in missing transport reaction frequently led to prediction errors. As an example, the *Salmonella* mutant SL1344 *yaeC metA* was auxotrophic for methionine and lacked the only annotated methionine transporter. However, this mutant could still grow in media containing µM

methionine quantities, suggesting efficient transport capabilities. Indeed, literature mining indicated at least one so far unidentified methionine transport system in *Salmonella* [112]. Possible candidates for this methionine uptake system include 18 annotated or putative amino acid transporters [112, 114]. Experimental analysis of these candidates could lead to the identification of the missing methionine transporter. Besides obvious knowledge gaps of non-annotated reactions, additional poorly known metabolic reactions are also lacking in the reconstruction. Some of these knowledge gaps can be identified by comparing *in silico* predictions with experimental data. For example, a gap of knowledge was indicated by the below average growth prediction accuracy on sulfur compounds (40%). This suggests that further research could focus on metabolic pathways involving *Salmonella* sulfur metabolism to identify missing enzyme annotations and reactions to bridge this knowledge gap.

3.2 A quantitative model of *Salmonella* metabolism during infection

Metabolism depends on nutrient availability. Any comprehensive qualitative understanding of *Salmonella* metabolism during infection thus requires data on *in vivo* nutrient availability.

Literature data suggests that *Salmonella* has access to a range of metabolites in varying quantities during infection. In different studies, the carbon sources glucose, gluconate and fatty acids were shown to be available in the SCV [69, 115, 116]. So far it is not known in what quantities these nutrients are available, and if these nutrients are the sole main carbon sources available to *Salmonella* during infection. Besides carbon sources, *Salmonella* has some access to a wide range of amino acids [48, 49, 66, 67, 117, 118], purines [67], pyrimidines [48, 49] and cofactors / cofactor precursors [72, 119] in its *in vivo* niche.

To determine the quantitative nutrient availability in the SCV, we tested the replication rate of multiple nutrient utilization and auxotrophic *Salmonella* mutants *in vivo*. Our experimental data and literature data evaluation suggested that at least 45 metabolites are available to *Salmonella in vivo.* Qualitative data for *in vivo* availability of 64 metabolites was already included by Raghunatan *et al.* into a *Salmonella* metabolic reconstruction, based on diverse but not cited literature data [100]. From these 64 metabolites, we experimentally confirmed availability of 29 nutrients. 12 nutrients only included in the

model of Raghunatan *et al.* had at most a minor relevance according to our experimental data and literature resources, including sialic acid, allantoin, gluconate, ethanolamine and propanediol [70, 72]. Since in Raghunathan *et al.* the corresponding literature sources were not specified, it is not possible to evaluate why specific nutrients were included into their model and if they limited the nutrient availability scope to a typhoid fever model [100]. A large number of the 35 nutrients either not tested by us or without detectable contribution are likely present during enteritis when *Salmonella* resides in the gut lumen, but not in typhoid fever (e.g., ethanolamine, dimethyl sulfoxide, cellobiose, fructose, fucose / arabinose, galactitol and galactonate [64, 70]).

Another explanation for discrepancies in nutrient availability might be that Raghunatan *et al.* included results of indirect, misleading experimental data into their model. We demonstrated that nutrient availability can be determined by replication rates of auxotrophic mutants and transporter mutants that are unable to take up and metabolize specific nutrients (see 2.2). Besides this approach, nutrient availability was frequently assumed by numerous indirect methods. For example, based on phenotypes of catabolic pathway mutants it was often inferred that the respective nutrient is an important part of the bacteria´s nutrition (e.g., fatty acid availability assumed based on isocitrate lyases Δ*icl1 Δicl2* phenotypes of *M. tuberculosis* [120]). However, mutations in nutrient catabolism can also lead to an accumulation of toxic compounds, hence these mutant phenotypes are not informative about quantitative nutrient availability [121]. Nutrient availability was also determined in literature based on transcriptome or proteome data. For example, gluconate was appointed to be one of the main carbon sources during infection, based on an upregulation of gluconate transporters in *Salmonella* infecting macrophages [116]. In comparison to this, Liu *et al*. demonstrated that transcriptome data can be misleading. They detected in *E. coli* growing on different carbon sources with decreasing replication rates an expanding transcription pattern, resulting in a systematic expression of catabolic pathways for compounds that were not present in the media [122]. Thus, expression of a certain transporter does not necessarily imply presence of its corresponding nutrient. Moreover, cell culture models might differ dramatically from *in vivo* conditions, hence results based on *in vitro* experiments might not reflect nutrient availability *in vivo* [123].

The metabolic model of Raghunatan *et al.* allowed with its qualitative nutrient availability a general *in vivo* phenotype prediction of “growth” versus “no growth”, leading to a correct prediction of 44 from 55 qualitative virulence phenotypes [100]. However, it prohibited precise growth rate determinations of

mutant phenotypes due to lack of quantitative nutrient exchange fluxes. Since we were interested in a comprehensive understanding of *Salmonella in vivo* metabolism, we needed to integrate quantitative nutrient availability into our model. To determine quantitative flux estimates, we evaluated experimental *in vivo* data of *Salmonella* nutrient utilization defect and auxotrophic mutants, and applied knowledge over biomass requirements and nutrient utilization efficiency. To further improve prediction accuracy, we limited 242 enzymatic reaction fluxes to their *in vivo* feasible range based on *ex vivo* protein abundance and enzyme turnover numbers ([124, 125], own unpublished data). We adapted biomass requirements to *Salmonella in vivo* phenotypes reported in literature for a more accurate representation of minimal requirements during infection. Based on all these data, we transformed our reconstruction of *Salmonella in vitro* metabolism (STMv1.0) into a model of *Salmonella in vivo* metabolism (STMv1.1). This model was extensively supported by large-scale experimental data including proteome data (see 2.3) and 799 *in vivo* phenotypes (see 2.2).

The model STMv1.1 allowed a comprehensive and quantitative analysis of *Salmonella* metabolism during infection. The evaluation of quantitative nutrient uptake fluxes revealed a paradoxical situation of a qualitatively nutrient rich, but quantitatively nutrient poor *in vivo* environment. We propose that this nutrient shortage leads to a limitation of *Salmonella in vivo* replication to the observed 6 h per division [64]. In line with this, any further diminishment of nutrient supply led to a reduced *Salmonella* growth rate.

The SCV is a nutrient poor environment, which is likely a special case for intracellular environments inhabited by pathogens. In comparison to this, other host compartments like blood / cerebrospinal fluid and the cytosol are known to be nutrient rich compartments [126]. This high quantity of nutrients is reflected by the rapid growth of diverse pathogens in these niches. For example, the cytosolic pathogen *Shigella* replicates with a generation time of 30 minutes in infected human cells, approaching growth rates observed in rich culture media [127-130]. The pathogen *Neisseria meningitis* on the other hand causes a rapid disease progression by fast replication in blood and / or cerebrospinal fluid, reaching loads of 10^9 bacteria per ml in these compartments ([131, 132], reviewed in [133]). These findings suggests rich, non-limiting nutrient supply at least for pathogens residing in compartments like blood / cerebrospinal fluid and cytoplasm, as opposed to a nutrient limitation for *Salmonella* and maybe also for other vacuolar pathogens. The vacuolar pathogen *M. tuberculosis* has a very low growth rate *in vivo* with generation times of one to four days, but a similarly slow replication occurs also in rich media *in vitro*

[134, 135], suggesting that factors unrelated to nutrition might be limiting. Thus simple conclusions of a general nutrient limitation for vacuolar pathogens cannot be drawn at present. The generation of quantitative *in vivo* models for other pathogens might help to unravel nutrient supply in various host environments.

Taken together, our model of *Salmonella in vivo* metabolism is the first that describes the quantitative *in vivo* nutritional landscape of a pathogen. It gives a comprehensive overview about *Salmonella* metabolism during systemic infection, including quantitative nutrient flux estimates and nutrient utilization during *Salmonella in vivo* growth. The results provide clear evidence of nutrient starvation as a limiting factor for *Salmonella* systemic virulence.

3.3 Analysis of robustness of *Salmonella in vivo* metabolism

Salmonella metabolism shows a high resilience against perturbations. This is demonstrated by the small number of metabolic genes that are essential for virulence [64]. Furthermore, some 75 % of over 800 *in vivo* phenotypes of metabolic genes described in literature had no detectable virulence contribution [72, 73]. This is similar to other pathogens, where a high percentage of genes does not contribute to virulence [72, 73, 136-142].

We wanted to analyze which factors contribute to the observed resilience of the *Salmonella* genome against perturbations, focusing on *in vivo* metabolism as a suitable subsystem [143]. Using our *in vivo* model of infection (see 2.2), we were interested in genes with undetectable phenotypes during infection and wanted to determine if they were either truly dispensable, dispensable due to redundancy or if they contributed to virulence in a marginal but maybe cumulative manner (“marginal benefit” hypothesis [144]).

It was already observed in diverse experimental and *in silico* based studies on various microbes, that the majority of metabolic genes is dispensable and inactive in a single environment [84, 108, 145]. Similarly, our analysis of *Salmonella in vivo* metabolism suggested that some 75% of *Salmonella* metabolic genes

were dispensable during infection. Further analysis indicated that the qualitatively nutrient rich host environment rendered the major part of *Salmonella*'s metabolic potential superfluous and inactive *in vivo*.

Apart from the *in vivo* nutrient availability, the presence of isofunctional enzymes ("intrinsic redundancy") or alternative pathways ("extrinsic" redundancy) are other factors with a high impact on metabolic robustness [108, 145-147]. Similarly to previous studies, our i*n silico* and experimental analyses of metabolism indicated that some 39% of all *in vivo* active enzymes are dispensable due to intrinsic (26%) or extrinsic (13%) redundancy.

Extensive gene inactivity and redundancy indicated that a major part of *Salmonella* metabolism is not needed in the typhoid fever model. Indeed, *in silico* analysis suggested that it is possible to simultaneously delete most genes with undetectable virulence phenotype, resulting in a minimal genome containing only 364 to 373 metabolic enzymes that still allowed for unimpaired *in vivo* growth. This suggests that the majority of *Salmonella* metabolic genes is truly dispensable for virulence in systemic infection, thus explaining the observed astonishing resilience against metabolic perturbations [64]. This result is in direct contradiction to the "marginal benefit" hypothesis of a small but accumulating fitness contributions of many enzymes with weak phenotypes [144]. On the other hand, *E. coli* strains lacking 15 to 22% of their genomes demonstrated equal growth rates to the wild type, which strongly speaks against a marginal but accumulative benefit of the removed 704 to 1081 genes [148, 149].

Another implication of this finding is that the qualitatively nutrient rich host environment allows for a massively reduced *Salmonella* genome, reminiscent of small genomes occurring in some pathogens such as *Mycoplasma* [150]. Consistent with this, previous *in silico* analyses of other microorganisms suggested that only some 245 genes and some 300 metabolic reactions are needed for growth in nutrient rich environments [151, 152].

In contrast to the core backbone of some 364 to 373 *Salmonella* genes needed for *in vivo* replication, our *ex vivo* proteome data and previous estimates indicated some 450 to 700 metabolic enzymes are actually expressed in the *Salmonella* typhoid fever model ([64], own unpublished data). This could be explained by (i) suboptimal regulation of inactive reactions (e.g. enzyme expression of the non-functional histidine synthesis pathway [66], or processes that were reported to contribute to virulence like (ii) high-

affinity uptake reactions which are neglected in FBA calculations if less cost intensive reactions are present (e.g. zinc uptake [153]) or (iii) reactions that are not included in the biomass function (e.g., reactive oxygen species (ROS) detoxification [154]). Thus, the actual minimal *Salmonella in vivo* metabolism might include more than the predicted backbone of some 364 to 373 enzymes. This could be analyzed by an experimental genome reduction approach.

3.4 Development of a method for large-scale gene inactivation in *Salmonella*

Salmonella in silico and *ex vivo* proteome analysis suggested that some 364 to 700 metabolic genes are needed for replication *in vivo,* which would be in the range of predictions for a minimal self-sufficient gene set in a nutrient rich environment [155, 156]. Validation of this hypothesis requires extensive genome-scale gene inactivation.

Bacteria with a reduced genome are occurring in nature. These bacteria were subject to different degrees of reductive evolution that led to the inactivation of a few genes (*S*. Typhi) to massive genome reduction to a minimal set of some hundred functional genes (*Buchnera*, *Mycoplasma*) ([150, 157, 158], reviewed in [159]).

Bacteria with reduced genome sizes can also be artificially generated with different experimental approaches, including synthesis of artificial chromosomes [160], targeted removal of large clusters of dispensable genes [148, 161, 162] or random mutagenesis [163-165]. Artificial synthesis and combination of all essential genes into a transplantable genome is a powerful approach to obtain a minimal genome [160]. However, this method requires enormous resources and is currently feasible only for small genomes such as *Mycoplasma* (reviewed in [166]). Targeted gene deletions on the other hand lead to initial fast removal of large dispensable gene clusters, but this approach becomes more and more time consuming due to (i) a steady decrease in gene cluster size and (ii) experimental setbacks because so far non-identified essential or synthetically lethal genes might have been inadvertently deleted. These might be some of the main reasons why no major progress for targeted gene deletion projects was reported within the last years [148, 149, 162].

In comparison to this, random mutagenesis approaches have the benefit that mutations occur without time consuming experimental setbacks, since surviving progeny accumulates inactivating mutations only in non-essential genes. Furthermore, these approaches are unbiased and allow multiple simultaneous experiments, improving the information content of the approach. So far, multiple random mutagenesis approaches were done for *S*. Typhimurium and *E. coli* [163-165], but none possessed a sufficiently high mutagenesis rate for large-scale mutagenesis.

We wanted to experimentally test our hypothesis of a largely dispensable *Salmonella* metabolism by using an unbiased large-scale random mutagenesis approach to inactivate as many genes as possible in the *Salmonella* genome. We therefore generated mutator lines through deleting the anti-mutator genes *dnaQ* and *mutS,* which caused an increase of the mutation rate by about 6.000 times compared to the wild type and by a factor of 20 to 26 compared to previous approaches [163, 165, 167]. We passaged 20 separate mutator lines in two different media conditions, complemented these strains with a *dnaQ mutS* expression plasmid and analyzed the stabilized lines via deep sequencing. Each line accumulated some 2700 to 3900 mutations, which caused inactivation of some 400 to 600 genes per line.

In silico analysis of *Salmonella in vivo* metabolism predicted that the majority of metabolic genes could be removed without impairing growth. Earlier studies showed that mutagenesis in combination with single cell bottlenecks (high genetic drift) led to a continuous growth rate decrease, an effect also known as Muller's ratchet [163, 164, 168]. To counter-select a growth rate decrease and to prevent detrimental effects of bottlenecks, some 10^4 to 10^5 CFU per line were passaged each day. Nonetheless, all lines showed a decreased growth rate compared to the wild type. This could be caused by the mutation of genes contributing to growth. On the other hand, it could also be caused by non-specific effects like (i) higher protein turnover caused by mutated, less stable enzymes, (ii) higher permeability of the inner membrane through integration of mutated membrane proteins or (iii) disturbed expression regulation, leading to a massive protein overexpression. Furthermore, mutations leading to frameshifts or premature stop codons in the coding sequence of enzymes can result in shortened peptides without enzymatic activity. Expression of these peptide fragments represents a futile energy expenditure, decreasing the overall fitness of the mutant strain.

Further passaging of the complemented strains in larger numbers than the aforementioned 10^4 to 10^5 CFU might allow to select for faster growth through suppressor mutations. Fitness increase was e.g. demonstrated for *mutS* strains in a previous approach after passaging without population bottlenecks

[164]. An example of suppressor mutations would be the loss of mutated genes that encode inactive peptide fragments, since this would prevent expenditure of futile energetic costs or the formation of toxic protein aggregates [169]. Another example for suppressor mutations that can lead to faster growth is the overexpression of chaperones. Maisnier-Patin *et al.* detected a two- to threefold overexpression of the chaperones DnaK and GroEL in mutated lines indicating a detrimental impact of protein aggregation caused by mutagenesis. Additional overexpression of the chaperones GroEL further increased fitness, supporting the strong impact of misfolded proteins [165]. In comparison to this, preliminary proteome analysis of the 20 *dnaQ mutS* mutator lines passaged in our approach indicated a higher protein concentration of GroEL / GroES and DnaK by an average of 1.6 ± 0.7 compared to the wild type.

All in all, we were able to show that deletion of the anti-mutator genes *dnaQ* and *mutS* can be used for unbiased large-scale mutagenesis. Furthermore, our results indicate that hundreds of mutations can be tolerated by individual *Salmonella* strains. In total about two-thirds of all *Salmonella* genes were inactivated at least once. It can be expected that further passaging of the uncomplemented mutator lines would lead to mutagenesis saturation and minimal genomes.

3.5 Conclusion

In this thesis, a quantitative and comprehensive understanding of *Salmonella* metabolism was achieved through a combination of *in silico*, *in vitro* and *in vivo* approaches. The resulting genome-scale model is the first metabolic *in vivo* model of any pathogen that contains quantitative nutrient flux estimates. Quantitative predictions indicated the paradoxical situation of *Salmonella* residing in a qualitatively nutrient rich, but quantitatively nutrient poor environment. This leads to a limitation of *Salmonella in vivo* replication rate but contributes to robustness of *Salmonella* metabolism. Furthermore, *in silico* and *in vitro* analysis of gene essentiality showed, that the majority of metabolic genes is truly dispensable in a defined environment, which largely explains *Salmonella* metabolic robustness.

In summary these findings indicate that metabolism is crucial for *Salmonella in vivo* growth, disease progression and development of novel control strategies. The general approach of metabolism analysis presented in this thesis might be applicable to other important human pathogens.

4. Outlook

4.1 *In vivo* metabolism models for other pathogens

In this thesis, the construction of and applications for a metabolic *Salmonella in vivo* model were presented. The outlined steps can principally be applied to construct metabolic *in vivo* models for any other pathogen. This would for example reveal if nutrient limitation is specific for *Salmonella* or common for intracellular pathogens residing in phagosomes. However, successful development of such comprehensive metabolism models would require (i) the availability of a high quality genome annotation as basis for the metabolic reconstruction, (ii) availability of an appropriate small animal model system and (iii) methods for genetic manipulation of the pathogen.

Already published experimental data or metabolic reconstructions could be used for further analysis, but obviously some of these conditions are challenging for important pathogens such as *Mycobacterium tuberculosis*.

4.2 Analysis of *in vivo* heterogeneity

Salmonella infection shows heterogeneous distribution for growth rate and anatomical localization. Whereas *Salmonella* is homogeneously distributed in liver [27], in spleen *Salmonella* is found in the red pulp and the bordering marginal zone, but not in the white pulp [170]. Importantly, heterogeneous *Salmonella in vivo* growth rates were observed [171, 172], which should have important consequences for differential metabolic activities in the various *Salmonella* subpopulations. FBA predictions of our model are based on constant and average replication rates in a constant environment, which obviously is

an oversimplification for heterogeneous *Salmonella* populations. For incorporation of *in vivo* heterogeneity into the *in vivo* metabolism model, *Salmonella* subpopulations with different replication rates will need to be separately analyzed as a basis for multiple different models that could aid to analyze the infection process in more detail.

4.3 The generation of minimal genome strains

Data presented in chapter 2.4 showed that unbiased large-scale random mutagenesis can be used for massive gene inactivation in *Salmonella*. For both "natural" minimal genome strains as well as for bacteria that have been subject to a targeted gene deletion approach, the additional application of random mutagenesis could offer major benefits. Further gene inactivation via random mutagenesis could lead to a fast and saturated profile of essential genes without time consuming experimental setbacks, since surviving progeny accumulates inactivating mutations only in non-essential genes. The inactivation of anti-mutator genes like *dnaQ* and *mutS* could be used for the generation of a minimal *in vitro* genome for a range of bacteria. In contrast to this, *in vivo* experiments would be hampered by a constant and high mutagenesis rate, causing dramatic virulence attenuation *in vivo*. For an *in vivo* minimal genome, transient mutagenesis only *in vitro* but not *in vivo* would be required. Exposure to UV light could represent a suitable random mutagenesis method for such applications ([173, 174], reviewed in [175]).

5. References

1. Thiele, I., et al., *A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2.* BMC systems biology, 2011. **5**: p. 8.
2. Madigan, M.T., J.M. Martinko, and T.D. Brock, *Brock biology of microorganisms*. 11th ed2006, Upper Saddle River, NJ: Pearson Prentice Hall. p.
3. Groisman, E.A. and H. Ochman, *How Salmonella became a pathogen.* Trends in microbiology, 1997. **5**(9): p. 343-9.
4. Lan, R., P.R. Reeves, and S. Octavia, *Population structure, origins and evolution of major Salmonella enterica clones.* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2009. **9**(5): p. 996-1005.
5. Herrera-Leon, S., et al., *Molecular characterization of a new serovar of Salmonella bongori 13,22:z39:- isolated from a lizard.* Research in microbiology, 2005. **156**(4): p. 597-602.
6. Edwards, P.R. and F. Kauffmann, *A simplification of the Kauffmann-White schema.* American journal of clinical pathology, 1952. **22**(7): p. 692-7.
7. Grimont, P. and F.-X. Weill, *Antigenic formulae of the Salmonella serovars, 9th edition.* World Health Organization Collaborating Centre for Reference and Research on Salmonella, Institute Pasteur, Paris, France, 2007.
8. Richardson, E.J., et al., *Genome sequences of Salmonella enterica serovar typhimurium, Choleraesuis, Dublin, and Gallinarum strains of well- defined virulence in food-producing animals.* Journal of bacteriology, 2011. **193**(12): p. 3162-3.
9. Stevens, M.P., T.J. Humphrey, and D.J. Maskell, *Molecular insights into farm animal and zoonotic Salmonella infections.* Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2009. **364**(1530): p. 2709-23.
10. Majowicz, S.E., et al., *The global burden of nontyphoidal Salmonella gastroenteritis.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2010. **50**(6): p. 882-9.
11. Crump, J.A., S.P. Luby, and E.D. Mintz, *The global burden of typhoid fever.* Bulletin of the World Health Organization, 2004. **82**(5): p. 346-53.
12. Coburn, B., G.A. Grassl, and B.B. Finlay, *Salmonella, the host and disease: a brief review.* Immunology and cell biology, 2007. **85**(2): p. 112-8.
13. Santos, R.L., et al., *Animal models of Salmonella infections: enteritis versus typhoid fever.* Microbes and infection / Institut Pasteur, 2001. **3**(14-15): p. 1335-44.
14. Bhan, M.K., R. Bahl, and S. Bhatnagar, *Typhoid and paratyphoid fever.* Lancet, 2005. **366**(9487): p. 749-62.
15. Watson, K.G. and D.W. Holden, *Dynamics of growth and dissemination of Salmonella in vivo.* Cellular microbiology, 2010. **12**(10): p. 1389-97.
16. Mason, W.P., *"Typhoid Mary".* Science, 1909. **30**(760): p. 117-8.
17. *Mary Mallon (Typhoid Mary).* American journal of public health and the nation's health, 1939. **29**(1): p. 66-8.
18. Gonzalez-Escobedo, G., J.M. Marshall, and J.S. Gunn, *Chronic and acute infection of the gall bladder by Salmonella Typhi: understanding the carrier state.* Nature reviews. Microbiology, 2011. **9**(1): p. 9-14.
19. Morpeth, S.C., H.O. Ramadhani, and J.A. Crump, *Invasive non-Typhi Salmonella disease in Africa.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2009. **49**(4): p. 606-11.

20. Tsolis, R.M., et al., *How to become a top model: impact of animal experimentation on human Salmonella disease research.* Infection and immunity, 2011. **79**(5): p. 1806-14.
21. Song, J., et al., *A mouse model for the human pathogen Salmonella typhi.* Cell host & microbe, 2010. **8**(4): p. 369-76.
22. Firoz Mian, M., et al., *Humanized mice are susceptible to Salmonella typhi infection.* Cellular & molecular immunology, 2011. **8**(1): p. 83-7.
23. Gunshin, H., et al., *Cloning and characterization of a mammalian proton-coupled metal-ion transporter.* Nature, 1997. **388**(6641): p. 482-8.
24. Goswami, T., et al., *Natural-resistance-associated macrophage protein 1 is an H+/bivalent cation antiporter.* The Biochemical journal, 2001. **354**(Pt 3): p. 511-9.
25. Belouchi, A., et al., *The macrophage-specific membrane protein Nramp controlling natural resistance to infections in mice has homologues expressed in the root system of plants.* Plant molecular biology, 1995. **29**(6): p. 1181-96.
26. Nakoneczna, I. and H.S. Hsu, *Histopathological study of protective immunity against murine salmonellosis induced by killed vaccine.* Infection and immunity, 1983. **39**(1): p. 423-30.
27. Richter-Dahlfors, A., A.M. Buchan, and B.B. Finlay, *Murine salmonellosis studied by confocal microscopy: Salmonella typhimurium resides intracellularly inside macrophages and exerts a cytotoxic effect on phagocytes in vivo.* The Journal of experimental medicine, 1997. **186**(4): p. 569-80.
28. Mert, A., et al., *Typhoid fever as a rare cause of hepatic, splenic, and bone marrow granulomas.* Internal medicine, 2004. **43**(5): p. 436-9.
29. Galan, J.E. and R. Curtiss, 3rd, *Virulence and vaccine potential of phoP mutants of Salmonella typhimurium.* Microbial pathogenesis, 1989. **6**(6): p. 433-43.
30. Gorden, J. and P.L. Small, *Acid resistance in enteric bacteria.* Infection and immunity, 1993. **61**(1): p. 364-7.
31. Carter, P.B. and F.M. Collins, *The route of enteric infection in normal mice.* The Journal of experimental medicine, 1974. **139**(5): p. 1189-203.
32. Francis, C.L., M.N. Starnbach, and S. Falkow, *Morphological and cytoskeletal changes in epithelial cells occur immediately upon interaction with Salmonella typhimurium grown under low-oxygen conditions.* Molecular microbiology, 1992. **6**(21): p. 3077-87.
33. Jones, B.D., N. Ghori, and S. Falkow, *Salmonella typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches.* The Journal of experimental medicine, 1994. **180**(1): p. 15-23.
34. Francis, C.L., et al., *Ruffles induced by Salmonella and other stimuli direct macropinocytosis of bacteria.* Nature, 1993. **364**(6438): p. 639-42.
35. Niess, J.H., et al., *CX3CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance.* Science, 2005. **307**(5707): p. 254-8.
36. Vazquez-Torres, A., et al., *Extraintestinal dissemination of Salmonella by CD18-expressing phagocytes.* Nature, 1999. **401**(6755): p. 804-8.
37. Voedisch, S., et al., *Mesenteric lymph nodes confine dendritic cell-mediated dissemination of Salmonella enterica serovar Typhimurium and limit systemic disease in mice.* Infection and immunity, 2009. **77**(8): p. 3170-80.
38. Monack, D.M., D.M. Bouley, and S. Falkow, *Salmonella typhimurium persists within macrophages in the mesenteric lymph nodes of chronically infected Nramp1+/+ mice and can be reactivated by IFNgamma neutralization.* The Journal of experimental medicine, 2004. **199**(2): p. 231-41.

39. Menendez, A., et al., *Salmonella infection of gallbladder epithelial cells drives local inflammation and injury in a model of acute typhoid fever.* The Journal of infectious diseases, 2009. **200**(11): p. 1703-13.
40. Alpuche-Aranda, C.M., et al., *Salmonella stimulate macrophage macropinocytosis and persist within spacious phagosomes.* The Journal of experimental medicine, 1994. **179**(2): p. 601-8.
41. Takeuchi, A., *Electron microscope studies of experimental Salmonella infection. I. Penetration into the intestinal epithelium by Salmonella typhimurium.* The American journal of pathology, 1967. **50**(1): p. 109-36.
42. Garcia-del Portillo, F. and B.B. Finlay, *Targeting of Salmonella typhimurium to vesicles containing lysosomal membrane glycoproteins bypasses compartments with mannose 6-phosphate receptors.* The Journal of cell biology, 1995. **129**(1): p. 81-97.
43. Buchmeier, N.A. and F. Heffron, *Inhibition of macrophage phagosome-lysosome fusion by Salmonella typhimurium.* Infection and immunity, 1991. **59**(7): p. 2232-8.
44. Rathman, M., M.D. Sjaastad, and S. Falkow, *Acidification of phagosomes containing Salmonella typhimurium in murine macrophages.* Infection and immunity, 1996. **64**(7): p. 2765-73.
45. Drecktrah, D., et al., *Salmonella trafficking is defined by continuous dynamic interactions with the endolysosomal system.* Traffic, 2007. **8**(3): p. 212-25.
46. Oh, Y.K., et al., *Rapid and complete fusion of macrophage lysosomes with phagosomes containing Salmonella typhimurium.* Infection and immunity, 1996. **64**(9): p. 3877-83.
47. Kinchen, J.M. and K.S. Ravichandran, *Phagosome maturation: going through the acid test.* Nature reviews. Molecular cell biology, 2008. **9**(10): p. 781-95.
48. Leung, K.Y. and B.B. Finlay, *Intracellular replication is essential for the virulence of Salmonella typhimurium.* Proceedings of the National Academy of Sciences of the United States of America, 1991. **88**(24): p. 11470-4.
49. Fields, P.I., et al., *Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent.* Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(14): p. 5189-93.
50. Kuhle, V., G.L. Abrahams, and M. Hensel, *Intracellular Salmonella enterica redirect exocytic transport processes in a Salmonella pathogenicity island 2-dependent manner.* Traffic, 2006. **7**(6): p. 716-30.
51. Mills, D.M., V. Bajaj, and C.A. Lee, *A 40 kb chromosomal fragment encoding Salmonella typhimurium invasion genes is absent from the corresponding region of the Escherichia coli K-12 chromosome.* Molecular microbiology, 1995. **15**(4): p. 749-59.
52. Barthel, M., et al., *Pretreatment of mice with streptomycin provides a Salmonella enterica serovar Typhimurium colitis model that allows analysis of both pathogen and host.* Infection and immunity, 2003. **71**(5): p. 2839-58.
53. Galan, J.E. and D. Zhou, *Striking a balance: modulation of the actin cytoskeleton by Salmonella.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(16): p. 8754-61.
54. Shea, J.E., et al., *Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(6): p. 2593-7.
55. Miller, S.I., A.M. Kukral, and J.J. Mekalanos, *A two-component regulatory system (phoP phoQ) controls Salmonella typhimurium virulence.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(13): p. 5054-8.
56. Fields, P.I., E.A. Groisman, and F. Heffron, *A Salmonella locus that controls resistance to microbicidal proteins from phagocytic cells.* Science, 1989. **243**(4894 Pt 1): p. 1059-62.

57. Bader, M.W., et al., *Recognition of antimicrobial peptides by a bacterial sensor kinase.* Cell, 2005. **122**(3): p. 461-72.
58. Prost, L.R., et al., *Activation of the bacterial sensor kinase PhoQ by acidic pH.* Molecular cell, 2007. **26**(2): p. 165-74.
59. Monsieurs, P., et al., *Comparison of the PhoPQ regulon in Escherichia coli and Salmonella typhimurium.* Journal of molecular evolution, 2005. **60**(4): p. 462-74.
60. Guo, L., et al., *Regulation of lipid A modifications by Salmonella typhimurium virulence genes phoP-phoQ.* Science, 1997. **276**(5310): p. 250-3.
61. Golubeva, Y.A. and J.M. Slauch, *Salmonella enterica serovar Typhimurium periplasmic superoxide dismutase SodCI is a member of the PhoPQ regulon and is induced in macrophages.* Journal of bacteriology, 2006. **188**(22): p. 7853-61.
62. Hensel, M., et al., *Genes encoding putative effector proteins of the type III secretion system of Salmonella pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages.* Molecular microbiology, 1998. **30**(1): p. 163-74.
63. Abrahams, G.L. and M. Hensel, *Manipulating cellular transport and immune responses: dynamic interactions between intracellular Salmonella enterica and its host cells.* Cellular microbiology, 2006. **8**(5): p. 728-37.
64. Becker, D., et al., *Robust Salmonella metabolism limits possibilities for new antimicrobials.* Nature, 2006. **440**(7082): p. 303-7.
65. Bumann, D., *Pathogen proteomes during infection: A basis for infection research and novel control strategies.* Journal of proteomics, 2010. **73**(11): p. 2267-76.
66. Hoiseth, S.K. and B.A. Stocker, *Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines.* Nature, 1981. **291**(5812): p. 238-9.
67. McFarland, W.C. and B.A. Stocker, *Effect of different purine auxotrophic mutations on mouse-virulence of a Vi-positive strain of Salmonella dublin and of two strains of Salmonella typhimurium.* Microbial pathogenesis, 1987. **3**(2): p. 129-41.
68. Tchawa Yimga, M., et al., *Role of gluconeogenesis and the tricarboxylic acid cycle in the virulence of Salmonella enterica serovar Typhimurium in BALB/c mice.* Infection and immunity, 2006. **74**(2): p. 1130-40.
69. Bowden, S.D., et al., *Glucose and glycolysis are required for the successful infection of macrophages and mice by Salmonella enterica serovar typhimurium.* Infection and immunity, 2009. **77**(7): p. 3117-26.
70. Thiennimitr, P., et al., *Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(42): p. 17480-5.
71. Winter, S.E., et al., *Gut inflammation provides a respiratory electron acceptor for Salmonella.* Nature, 2010. **467**(7314): p. 426-9.
72. Chaudhuri, R.R., et al., *Comprehensive identification of Salmonella enterica serovar typhimurium genes required for infection of BALB/c mice.* PLoS pathogens, 2009. **5**(7): p. e1000529.
73. Santiviago, C.A., et al., *Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice.* PLoS pathogens, 2009. **5**(7): p. e1000477.
74. Hucka, M., et al., *The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models.* Bioinformatics, 2003. **19**(4): p. 524-31.
75. Fleming, R.M., I. Thiele, and H.P. Nasheuer, *Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to Escherichia coli.* Biophysical chemistry, 2009. **145**(2-3): p. 47-56.

76. Schellenberger, J., et al., *BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions.* BMC bioinformatics, 2010. **11**: p. 213.
77. Gutnick, D., et al., *Compounds which serve as the sole source of carbon or nitrogen for Salmonella typhimurium LT-2.* Journal of bacteriology, 1969. **100**(1): p. 215-9.
78. Tracy, B.S., K.K. Edwards, and A. Eisenstark, *Carbon and nitrogen substrate utilization by archival Salmonella typhimurium LT2 cells.* BMC evolutionary biology, 2002. **2**: p. 14.
79. Eisenreich, W., et al., *Carbon metabolism of intracellular bacterial pathogens and possible links to virulence.* Nature reviews. Microbiology, 2010. **8**(6): p. 401-12.
80. Orth, J.D., et al., *A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011.* Molecular systems biology, 2011. **7**: p. 535.
81. Feist, A.M., et al., *A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information.* Molecular systems biology, 2007. **3**: p. 121.
82. Schuetz, R., L. Kuepfer, and U. Sauer, *Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli.* Molecular systems biology, 2007. **3**: p. 119.
83. Edwards, J.S. and B.O. Palsson, *Metabolic flux balance analysis and the in silico analysis of Escherichia coli K-12 gene deletions.* BMC bioinformatics, 2000. **1**: p. 1.
84. Lewis, N.E., et al., *Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models.* Molecular systems biology, 2010. **6**: p. 390.
85. Orth, J.D., I. Thiele, and B.O. Palsson, *What is flux balance analysis?* Nature biotechnology, 2010. **28**(3): p. 245-8.
86. Liao, Y.C., et al., *An experimentally validated genome-scale metabolic reconstruction of Klebsiella pneumoniae MGH 78578, iYL1228.* Journal of bacteriology, 2011. **193**(7): p. 1710-7.
87. Bordbar, A., et al., *Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions.* Molecular systems biology, 2010. **6**: p. 422.
88. Jamshidi, N. and B.O. Palsson, *Investigating the metabolic capabilities of Mycobacterium tuberculosis H37Rv using the in silico strain iNJ661 and proposing alternative drug targets.* BMC systems biology, 2007. **1**: p. 26.
89. Oberhardt, M.A., et al., *Genome-scale metabolic network analysis of the opportunistic pathogen Pseudomonas aeruginosa PAO1.* Journal of bacteriology, 2008. **190**(8): p. 2790-803.
90. McClelland, M., et al., *Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.* Nature, 2001. **413**(6858): p. 852-6.
91. Ochman, H. and A.C. Wilson, *Evolution in bacteria: evidence for a universal substitution rate in cellular genomes.* Journal of molecular evolution, 1987. **26**(1-2): p. 74-86.
92. Zhu, M., et al., *Functions of the siderophore esterases IroD and IroE in iron-salmochelin utilization.* Microbiology, 2005. **151**(Pt 7): p. 2363-72.
93. Crouch, M.L., et al., *Biosynthesis and IroC-dependent export of the siderophore salmochelin are essential for virulence of Salmonella enterica serovar Typhimurium.* Molecular microbiology, 2008. **67**(5): p. 971-83.
94. Lin, H., et al., *In vitro characterization of salmochelin and enterobactin trilactone hydrolases IroD, IroE, and Fes.* Journal of the American Chemical Society, 2005. **127**(31): p. 11075-84.
95. Fischbach, M.A., et al., *In vitro characterization of IroB, a pathogen-associated C-glycosyltransferase.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(3): p. 571-6.
96. Hantke, K., et al., *Salmochelins, siderophores of Salmonella enterica and uropathogenic Escherichia coli strains, are recognized by the outer membrane receptor IroN.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(7): p. 3677-82.

97. Kroger, C. and T.M. Fuchs, *Characterization of the myo-inositol utilization island of Salmonella enterica serovar Typhimurium.* Journal of bacteriology, 2009. **191**(2): p. 545-54.
98. Kroger, C., J. Stolz, and T.M. Fuchs, *myo-Inositol transport by Salmonella enterica serovar Typhimurium.* Microbiology, 2010. **156**(Pt 1): p. 128-38.
99. Diaz, E., et al., *Biodegradation of aromatic compounds by Escherichia coli.* Microbiology and molecular biology reviews : MMBR, 2001. **65**(4): p. 523-69, table of contents.
100. Raghunathan, A., et al., *Constraint-based analysis of metabolic capacity of Salmonella typhimurium during host-pathogen interaction.* BMC systems biology, 2009. **3**: p. 38.
101. Oberhardt, M.A., B.O. Palsson, and J.A. Papin, *Applications of genome-scale metabolic reconstructions.* Molecular systems biology, 2009. **5**: p. 320.
102. Benedict, M.N., et al., *Genome-Scale Metabolic Reconstruction and Hypothesis Testing in the Methanogenic Archaeon Methanosarcina acetivorans C2A.* Journal of bacteriology, 2012. **194**(4): p. 855-65.
103. Charusanti, P., et al., *An experimentally-supported genome-scale metabolic network reconstruction for Yersinia pestis CO92.* BMC systems biology, 2011. **5**: p. 163.
104. AbuOun, M., et al., *Genome scale reconstruction of a Salmonella metabolic model: comparison of similarity and differences with a commensal Escherichia coli strain.* The Journal of biological chemistry, 2009. **284**(43): p. 29480-8.
105. Fang, K., et al., *Exploring the metabolic network of the epidemic pathogen Burkholderia cenocepacia J2315 via genome-scale reconstruction.* BMC systems biology, 2011. **5**: p. 83.
106. Deutscher, D., et al., *Multiple knockout analysis of genetic robustness in the yeast metabolic network.* Nature genetics, 2006. **38**(9): p. 993-8.
107. Suthers, P.F., A. Zomorrodi, and C.D. Maranas, *Genome-scale gene/reaction essentiality and synthetic lethality analysis.* Molecular systems biology, 2009. **5**: p. 301.
108. Papp, B., C. Pal, and L.D. Hurst, *Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast.* Nature, 2004. **429**(6992): p. 661-4.
109. MacDonald, S.J., G.H. Thomas, and A.E. Douglas, *Genetic and metabolic determinants of nutritional phenotype in an insect-bacterial symbiosis.* Molecular ecology, 2011. **20**(10): p. 2073-84.
110. Cosgriff, A.J. and A.J. Pittard, *A topological model for the general aromatic amino acid permease, AroP, of Escherichia coli.* Journal of bacteriology, 1997. **179**(10): p. 3317-23.
111. Linka, N. and A.P. Weber, *Intracellular metabolite transporters in plants.* Molecular plant, 2010. **3**(1): p. 21-53.
112. Ayling, P.D., T. Mojica-a, and T. Klopotowski, *Methionine transport in Salmonella typhimurium: evidence for at least one low-affinity transport system.* Journal of general microbiology, 1979. **114**(2): p. 227-46.
113. Johnson, D.A., et al., *High-throughput phenotypic characterization of Pseudomonas aeruginosa membrane transport genes.* PLoS genetics, 2008. **4**(10): p. e1000211.
114. Pruitt, K.D., et al., *NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy.* Nucleic acids research, 2012. **40**(Database issue): p. D130-5.
115. Fang, F.C., et al., *Isocitrate lyase (AceA) is required for Salmonella persistence but not for acute lethal infection in mice.* Infection and immunity, 2005. **73**(4): p. 2547-9.
116. Eriksson, S., et al., *Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica.* Molecular microbiology, 2003. **47**(1): p. 103-18.
117. Klose, K.E. and J.J. Mekalanos, *Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates Salmonella typhimurium virulence.* Infection and immunity, 1997. **65**(2): p. 587-96.

118. Das, P., et al., *Cationic amino acid transporters and Salmonella Typhimurium ArgT collectively regulate arginine availability towards intracellular Salmonella growth.* PloS one, 2010. **5**(12): p. e15466.
119. Bergthorsson, U. and J.R. Roth, *Natural isolates of Salmonella enterica serovar Dublin carry a single nadA missense mutation.* Journal of bacteriology, 2005. **187**(1): p. 400-3.
120. Munoz-Elias, E.J. and J.D. McKinney, *Mycobacterium tuberculosis isocitrate lyases 1 and 2 are jointly required for in vivo growth and virulence.* Nature medicine, 2005. **11**(6): p. 638-44.
121. Upton, A.M. and J.D. McKinney, *Role of the methylcitrate cycle in propionate metabolism and detoxification in Mycobacterium smegmatis.* Microbiology, 2007. **153**(Pt 12): p. 3973-82.
122. Liu, M., et al., *Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli.* The Journal of biological chemistry, 2005. **280**(16): p. 15921-7.
123. Sheppard, M., et al., *Dynamics of bacterial growth and distribution within the liver during Salmonella infection.* Cellular microbiology, 2003. **5**(9): p. 593-600.
124. Keseler, I.M., et al., *EcoCyc: a comprehensive database of Escherichia coli biology.* Nucleic acids research, 2011. **39**(Database issue): p. D583-90.
125. Scheer, M., et al., *BRENDA, the enzyme information system in 2011.* Nucleic acids research, 2011. **39**(Database issue): p. D670-6.
126. Wishart, D.S., et al., *HMDB: a knowledgebase for the human metabolome.* Nucleic acids research, 2009. **37**(Database issue): p. D603-10.
127. Goetz, M., et al., *Microinjection and growth of bacteria in the cytosol of mammalian host cells.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(21): p. 12221-6.
128. Lucchini, S., et al., *Transcriptional adaptation of Shigella flexneri during infection of macrophages and epithelial cells: insights into the strategies of a cytosolic bacterial pathogen.* Infection and immunity, 2005. **73**(1): p. 88-102.
129. Labrec, E.H., et al., *Epithelial Cell Penetration as an Essential Step in the Pathogenesis of Bacillary Dysentery.* Journal of bacteriology, 1964. **88**(5): p. 1503-18.
130. Gerber, D.F. and H.M. Watkins, *Growth of shigellae in monolayer tissue cultures.* Journal of bacteriology, 1961. **82**: p. 815-22.
131. Hackett, S.J., et al., *Meningococcal bacterial DNA load at presentation correlates with disease severity.* Archives of disease in childhood, 2002. **86**(1): p. 44-6.
132. Ovstebo, R., et al., *Use of robotized DNA isolation and real-time PCR to quantify and identify close correlation between levels of Neisseria meningitidis DNA and lipopolysaccharides in plasma and cerebrospinal fluid from patients with systemic meningococcal disease.* Journal of clinical microbiology, 2004. **42**(7): p. 2980-7.
133. Ferguson, L.E., et al., *Neisseria meningitidis: presentation, treatment, and prevention.* Journal of pediatric health care : official publication of National Association of Pediatric Nurse Associates & Practitioners, 2002. **16**(3): p. 119-24.
134. Gill, W.P., et al., *A replication clock for Mycobacterium tuberculosis.* Nature medicine, 2009. **15**(2): p. 211-4.
135. Armstrong, J.A. and P.D. Hart, *Response of cultured macrophages to Mycobacterium tuberculosis, with observations on fusion of lysosomes with phagosomes.* The Journal of experimental medicine, 1971. **134**(3 Pt 1): p. 713-40.
136. Hensel, M., et al., *Simultaneous identification of bacterial virulence genes by negative selection.* Science, 1995. **269**(5222): p. 400-3.
137. Kavermann, H., et al., *Identification and characterization of Helicobacter pylori genes essential for gastric colonization.* The Journal of experimental medicine, 2003. **197**(7): p. 813-22.

138. Maroncle, N., et al., *Identification of Klebsiella pneumoniae genes involved in intestinal colonization and adhesion using signature-tagged mutagenesis.* Infection and immunity, 2002. **70**(8): p. 4729-34.
139. Edelstein, P.H., et al., *Discovery of virulence genes of Legionella pneumophila by using signature tagged mutagenesis in a guinea pig pneumonia model.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(14): p. 8190-5.
140. Sun, Y.H., et al., *Functional genomics of Neisseria meningitidis pathogenesis.* Nature medicine, 2000. **6**(11): p. 1269-73.
141. Chiang, S.L. and J.J. Mekalanos, *Use of signature-tagged transposon mutagenesis to identify Vibrio cholerae genes critical for colonization.* Molecular microbiology, 1998. **27**(4): p. 797-805.
142. Camacho, L.R., et al., *Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis.* Molecular microbiology, 1999. **34**(2): p. 257-67.
143. Bumann, D., *System-level analysis of Salmonella metabolism during infection.* Current opinion in microbiology, 2009. **12**(5): p. 559-67.
144. Thatcher, J.W., J.M. Shaw, and W.J. Dickinson, *Marginal fitness contributions of nonessential genes in yeast.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(1): p. 253-7.
145. Blank, L.M., L. Kuepfer, and U. Sauer, *Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast.* Genome biology, 2005. **6**(6): p. R49.
146. Hartman, J.L.t., B. Garvik, and L. Hartwell, *Principles for the buffering of genetic variation.* Science, 2001. **291**(5506): p. 1001-4.
147. Gu, Z., et al., *Role of duplicate genes in genetic robustness against null mutations.* Nature, 2003. **421**(6918): p. 63-6.
148. Posfai, G., et al., *Emergent properties of reduced-genome Escherichia coli.* Science, 2006. **312**(5776): p. 1044-6.
149. Mizoguchi, H., et al., *Superpositioning of deletions promotes growth of Escherichia coli with a reduced genome.* DNA research : an international journal for rapid publication of reports on genes and genomes, 2008. **15**(5): p. 277-84.
150. Fraser, C.M., et al., *The minimal gene complement of Mycoplasma genitalium.* Science, 1995. **270**(5235): p. 397-403.
151. Pal, C., et al., *Chance and necessity in the evolution of minimal metabolic networks.* Nature, 2006. **440**(7084): p. 667-70.
152. Nishikawa, T., N. Gulbahce, and A.E. Motter, *Spontaneous reaction silencing in metabolic optimization.* PLoS computational biology, 2008. **4**(12): p. e1000236.
153. Ammendola, S., et al., *High-affinity Zn2+ uptake system ZnuABC is required for bacterial zinc homeostasis in intracellular environments and contributes to the virulence of Salmonella enterica.* Infection and immunity, 2007. **75**(12): p. 5867-76.
154. Uzzau, S., L. Bossi, and N. Figueroa-Bossi, *Differential accumulation of Salmonella[Cu, Zn] superoxide dismutases SodCI and SodCII in intracellular bacteria: correlation with their relative contribution to pathogenicity.* Molecular microbiology, 2002. **46**(1): p. 147-56.
155. Mushegian, A.R. and E.V. Koonin, *A minimal gene set for cellular life derived by comparison of complete bacterial genomes.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(19): p. 10268-73.
156. Koonin, E.V., *How many genes can make a cell: the minimal-gene-set concept.* Annual review of genomics and human genetics, 2000. **1**: p. 99-116.
157. Parkhill, J., et al., *Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.* Nature, 2001. **413**(6858): p. 848-52.

158. Shigenobu, S., et al., *Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS.* Nature, 2000. **407**(6800): p. 81-6.
159. Wixon, J., *Featured organism: reductive evolution in bacteria: Buchnera sp., Rickettsia prowazekii and Mycobacterium leprae.* Comparative and functional genomics, 2001. **2**(1): p. 44-8.
160. Gibson, D.G., et al., *Creation of a bacterial cell controlled by a chemically synthesized genome.* Science, 2010. **329**(5987): p. 52-6.
161. Kolisnychenko, V., et al., *Engineering a reduced Escherichia coli genome.* Genome research, 2002. **12**(4): p. 640-7.
162. Hashimoto, M., et al., *Cell size and nucleoid organization of engineered Escherichia coli cells with a reduced genome.* Molecular microbiology, 2005. **55**(1): p. 137-49.
163. Funchain, P., et al., *The consequences of growth of a mutator strain of Escherichia coli as measured by loss of function among multiple gene targets and loss of fitness.* Genetics, 2000. **154**(3): p. 959-70.
164. Nilsson, A.I., et al., *Bacterial genome size reduction by experimental evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(34): p. 12112-6.
165. Maisnier-Patin, S., et al., *Genomic buffering mitigates the effects of deleterious mutations in bacteria.* Nature genetics, 2005. **37**(12): p. 1376-9.
166. Jewett, M.C. and A.C. Forster, *Update on designing and building minimal cells.* Current opinion in biotechnology, 2010. **21**(5): p. 697-703.
167. Wielgoss, S., et al., *Mutation Rate Inferred From Synonymous Substitutions in a Long-Term Evolution Experiment With Escherichia coli.* G3, 2011. **1**(3): p. 183-186.
168. Felsenstein, J., *The evolutionary advantage of recombination.* Genetics, 1974. **78**(2): p. 737-56.
169. Kuo, C.H. and H. Ochman, *The extinction dynamics of bacterial pseudogenes.* PLoS genetics, 2010. **6**(8).
170. Salcedo, S.P., et al., *Intracellular replication of Salmonella typhimurium strains in specific subsets of splenic macrophages in vivo.* Cellular microbiology, 2001. **3**(9): p. 587-97.
171. Helaine, S., et al., *Dynamics of intracellular bacterial replication at the single cell level.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(8): p. 3746-51.
172. Grant, A.J., et al., *Modelling within-host spatiotemporal dynamics of invasive bacterial disease.* PLoS biology, 2008. **6**(4): p. e74.
173. Alcantara-Diaz, D., M. Brena-Valle, and J. Serment-Guerrero, *Divergent adaptation of Escherichia coli to cyclic ultraviolet light exposures.* Mutagenesis, 2004. **19**(5): p. 349-54.
174. Goldman, R.P. and M. Travisano, *Experimental evolution of ultraviolet radiation resistance in Escherichia coli.* Evolution; international journal of organic evolution, 2011. **65**(12): p. 3486-98.
175. Ikehata, H. and T. Ono, *The mechanisms of UV mutagenesis.* Journal of radiation research, 2011. **52**(2): p. 115-25.

6. Acknowledgements

First of all, I am especially grateful to Professor Dirk Bumann for letting me work on these interesting and multifaceted projects, for support and advice during my thesis and for always having an open door and time for discussions. I also thank Professor Dirk Bumann and Professor Christoph Dehio for correcting my thesis.

I would like to thank the members of my PhD committee, Professor Dirk Bumann, Professor Christoph Dehio and Professor Urs Jenal for their scientific advice during my thesis.

I also thank H. Maki for the supply of a *dnaE173* plasmid, M. Marinus for the *dam* plasmid and A. Böhm for supply of the strains BW20767 and BW21038 pLD54 used in different projects in my thesis.

I would like to thank all colleagues that helped me during my thesis. In specific, I am thankful to Beatrice Claudi and Alain Mazé for helping me getting started in our lab. I would like to thank Alain Mazé, Beatrice Claudi and Neil Burton for support in animal experiments, Alexander Schmidt, Erik Ahrné and Timo Glatter for support with proteome methods, Pablo Manfredi for help with OrthoMCL, MAQ and Perl, Mario Emmenlauer for help with VAAL and both of them for support in getting me started with Linux.

A big thanks to Julien Limenitakis, Olin Silander and Felix Steeb for helpful comments and advice after reading my thesis.

I thank past and present lab members for the pleasant atmosphere in our group. I also would like to thank Julien for the cookies in a desperate time of need.

A special thanks to my family (my mother, my brother Felix and my sister-in-law Alexandra) as well as to my friends for support, kindness and cheerfulness during my PhD thesis.

And most of all, I thank Kathrin for the past nine years and the life to come.

7. Supplemental information

7.1 Extensive *in vivo* resilience of persistent *Salmonella*

(Manuscript submitted)

Barat S, **Steeb B**, Bumann D

Summary:

Chronic infections with persistent pathogens are a major health problem, but *in vivo* models are largely lacking. We established a mouse model to identify genes needed for persistency in *Salmonella*. Our results indicate relaxed requirements for the persistence state of *Salmonella*. Out of 12 targets that were essential during acute infections, only five showed a phenotype in our persistency model. Amongst these five hits, only inactivation of unsaturated/cyclopropane fatty acid synthesis lead to a rapid clearance of persistent *Salmonella* during infection. *In silico* predictions did not indicate any other essential genes besides the fatty acid synthesis pathway, reflecting redundancy in providing all required precursors.

Statement of my work:

I contributed to this publication by predicting essential genes in the persistency model. For this, I used an *in vivo* model of *Salmonella* metabolism (STMv1.1, see 2.2). Furthermore, I generated and tested the virulence of SL1344 *gutQ yrbH*, which was then used as an additional target for the persistency model.

Extensive In Vivo Resilience of Persistent *Salmonella*

Short title: *Salmonella* persistence

Somedutta Barat, Benjamin Steeb, Dirk Bumann*

Focal Area Infection Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland

Corresponding author:

E-Mail: Dirk.Bumann@unibas.ch

Tel.: +41 (61) 267 23 82

Fax: +41 (61) 267 21 18

Abstract

Chronic infections caused by persistent pathogens represent an important health problem. Here, we establish a simple practical mouse *Salmonella* infection model for identifying bacterial maintenance functions that are essential for persistency. In this model, a substantial fraction of *Salmonella* survived even several days of treatment with a potent fluoroquinolone antibiotic indicating stringency of the model. Evaluation of twelve *Salmonella* defects revealed dramatically relaxed requirements for *Salmonella* during persistency as compared to acute infections. A defect in synthesis of unsaturated/cyclopropane fatty acids resulted in rapid *Salmonella* clearance suggesting that this pathway might contain suitable targets for antimicrobial chemotherapy of chronic infection.

Introduction

Persistent pathogens represent a major problem for control of infectious diseases [1]. Extensive drug tolerance of persisters to all available antimicrobials often leads to treatment failures and relapse. Persistent pathogens may adopt a non-replicating dormant stage with no requirement for macromolecular synthesis comprising most current antimicrobial targets [2]. As a consequence, such dormant stages are tolerant to most antibiotics. In addition, low metabolic activity during dormancy might minimize vulnerability to perturbation. In fact, it remains unclear if dormant persisters have any essential maintenance requirements for survival that could provide opportunities for eradication through antimicrobial chemotherapy.

Various in vitro models have been used as an approximation of chronic infection with dormant persisters. Data obtained with these models revealed differential perturbation effects depending on the particular model and the respective pathogen [3-8]. As an example, proton motive force-driven ATP synthesis has been shown to be essential for *Mycobacterium tuberculosis* survival in a hypoxia in vitro model [9]. Indeed, inhibition of ATP synthase accelerates mycobacterial eradication in patients [10]. On the other hand, diminishing ATP levels can actually promote *E. coli* in vitro persister formation [1]. Additional in vivo models could help to compare persister maintenance requirements under relevant conditions. However, except for *Mycobacterium tuberculosis*, practical in vivo persistency models are largely lacking.

Salmonella enterica can cause diarrhea or systemic disease called typhoid/paratyphoid fever. A substantial fraction of systemically infected individuals

develops asymptomatic chronic infection [11, 12]. In many cases, *Salmonella* persists in biofilms on gallstones but persisting *Salmonella* have also been detected in liver [13] and lymph nodes [14]. Surgical removal of gallstones is required for successful treatment of *Salmonella* in gallstones biofilms, while extended treatment with potent fluoroquinolone antibiotics is recommended for treatment of chronic *Salmonella* tissue colonization [15]. *Salmonella* physiology during persistency is largely unknown although some factors that might support chronic *Salmonella* survival have been identified [12].

In this study, we used a simple chronic mouse *Salmonella* infection model in which a substantial *Salmonella* subpopulation survived prolonged treatment with a fluoroquinolone antibiotic. In this stringent in vivo model, almost all tested *Salmonella* activities were dispensable confirming extensive resilience of persistent pathogens against perturbation. On the other hand, the data also revealed a few novel candidate targets that could be explored for their suitability to control chronic infections.

Results

Persistency model using *Salmonella purA ssaGH*

To generate a practical *Salmonella* persistency model, we constructed a *Salmonella* strain that survives but largely fails to proliferate in systemically infected mice. Specifically, we combined two mutations that had previously been shown to impair *Salmonella* in vivo growth: *purA* which blocks adenosine biosynthesis [16], and *ssaGH* which inactivates the SPI-2 (*Salmonella* pathogenicity island 2)-associated type three secretion system required for intracellular *Salmonella* growth and virulence [17]. Both *purA* and SPI-2 mutations have previously been shown to result in long-term persistence with minimal acute virulence, but our initial characterization revealed some initial in vivo proliferation of both individual mutants after i.v. administration (data not shown). In contrast, the double mutant *Salmonella purA ssaGH* had largely constant bacterial tissue loads in spleen and liver (Fig. 1A, B) suggesting limited net growth.

To determine the suitability of this model for evaluating antimicrobial targets during persistency, we treated infected mice with the antibiotic enrofloxacin. This antibiotic belongs to the fluoroquinolone class, which is uniquely effective against non-growing bacteria in vitro [18], and the treatment of choice for human persistent salmonellosis [15]. Enrofloxacin is also the most effective drug in the mouse typhoid fever where it diminishes *Salmonella* tissue loads to levels below detection within one day of treatment [19]. In our persistency model, enrofloxacin initially diminished spleen loads of *Salmonella purA ssaGH*, but within two days substantial surviving

subpopulation stabilized and remained clearly detectable during at least four days of treatment (Fig. 1A). Surprisingly, however, liver loads continuously decreased during prolonged treatment suggesting somewhat different *Salmonella* physiological states in the two host tissues. The substantial tolerance of *Salmonella purA ssaGH* in spleen against enrofloxacin indicated that our model offered a practical approach to study treatment failures during persistency. Enrofloxacin efficacy also provided a suitable benchmark for potential new *Salmonella* persistency targets.

Salmonella defects with minor persistency phenotypes

Only a small number of *Salmonella* genes are absolutely essential for *Salmonella* survival and growth in host tissues during acute salmonellosis [20]. Some of these genes might also be relevant for *Salmonella* persistency. To test this hypothesis we transduced 12 mutations into the parental *Salmonella purA ssaGH* strain. Mouse infections revealed that most tested genes had no impact on *Salmonella* persistency in our model despite their crucial importance during acute infections (Fig. 2).

As an example, *ubiC* encoding chorismate lyase is required for ubiquinone biosynthesis. During acute infection, *Salmonella ubiC* is completely cleared from infected mice within one day indicating absolute essentiality [20]. In striking contrast, *Salmonella purA ssaGH ubiC* survived at high levels indicating dispensability of ubiquinone-mediated oxidative respiration during persistency. Similarly, functional ATPase is essential for acute *Salmonella* virulence [21], but we found it to be fully dispensable during persistency. This was in striking contrast to various *Mycobacterium*

tuberculosis models that suggest ATPase to be a particularly attractive target for this pathogen [9, 10]. Another case with strikingly different relevance in acute [22] vs. persistent *Salmonella* infections was *recA* involved in DNA repair. More expectedly, *trxA* encoding a thioredoxin essential for SPI-2 function [23], had no detectable role in *Salmonella purA ssaGH* presumably because SPI-2 was already inactive in this strain.

Polyphosphate biosynthesis or fatty acid degradation were known to be largely dispensable during acute infection but had some role in other chronic *Salmonella* infection models [24, 25]. However, in our stringent model both activities had weak effects indicating their dispensability for persistency. All these negative results suggested a severely limited number of suitable targets for control of persistent *Salmonella* infections.

Salmonella defects with moderate persistency phenotypes

In contrast to all these cases, two mutations, *asd* and *gutQ yrbH*, showed moderate phenotypes in our model (Fig. 3). *asd* encoding aspartate semialdehyde dehydrogenase is required for biosynthesis of the cell-wall peptidoglycan component diaminopimelic acid. A *Salmonella asd* strain spontaneously lyses in vitro and is completely cleared within one day from systemically infected mice [20]. However, *Salmonella purA ssaGH asd* was only partially cleared during the first day post infection which might reflect residual proliferation of some *Salmonella* and/or difficulties in establishing a suitable systemic niche [26]. Thereafter, this strain persisted at slowly declining levels in spleen. This could reflect non-essentiality of cell-wall synthesis for non-growing bacteria [18]. In

contrast, liver loads rapidly declined suggesting a substantial fraction of *Salmonella purA ssaGH* with active cell-wall turnover/growth in liver. This higher *Salmonella* activity in liver as compared to spleen was consistent with the differential effect of continuous enrofloxacin treatment on *Salmonella purA ssaGH* loads in these two tissues (Fig. 1). Similarly, *Salmonella purA ssaGH gutQ yrbH* that required supplementation with the lipopolysacharide precursor arabinose-5-phosphate to grow in vitro [27] and was highly attenuated during acute infections (our unpublished data), maintained high levels in spleen but was cleared from liver suggesting limited lipopolysaccharide demands during *Salmonella* persistency. Both genes thus were unsuitable as targets.

Two additional mutants had very severe colonization defects but still maintained stable small loads indicating non-essentiality for seven day persistence. *Salmonella purA ssaGH ribB* defective for 3,4-dihydroxy-2-butanone 4-phosphate synthase which is involved in riboflavin biosynthesis, was cleared within one day post infection to very low levels in both spleen and liver, but stabilized thereafter particularly in liver. This might reflect differential availability of host riboflavin supplementation in these two tissues. Importantly, these data showed that *Salmonella* with defective riboflavin biosynthesis can survive in vivo for extended periods. Another strain that was rapidly cleared from spleen had a defect in *iscS* encoding cysteine desulfurase involved in repair of iron-sulfur clusters and tRNA modification [28]. This mutant also dropped to very low loads in liver but still maintained detectable loads at seven days post infection.

β-ketoacyl-ACP synthase I essentiality for *Salmonella* persistency

Finally, there was a single mutant with a more promising phenotype (Fig. 4A). *Salmonella purA ssaGH fabB* defective for β-ketoacyl-ACP synthase I required for biosynthesis of unsaturated fatty acids and cyclopropane fatty acids, was progressively cleared from both liver and spleen. During clearance, residual *Salmonella purA ssaGH fabB* were recovered from mice mostly as small-colony variants. Withdrawal of fatty acid supplementation in vitro similarly enriched small-colony variants of this strain (Fig. 4B), suggesting that reduced growth and metabolism might enhance survival of this mutant when external fatty acids are unavailable. However, even small-colony variants were rapidly cleared from mouse tissues. Small-colony variants generated in vivo or in vitro quickly reverted to fast growth upon sub-culturing in presence of oleic acid supplementation.

To test the suitability of this target for antimicrobial chemotherapy, we administered thiolactomycin [29], a slow onset inhibitor of β-ketoacyl-ACP synthase I that is effective in mouse infection models with extracellular pathogens [30]. However, safe doses did not diminish *Salmonella purA ssaGH* loads in spleen (Fig. 4C). This could reflect the low target affinity of this inhibitor and/or poor delivery to *Salmonella* that persist intracellularly in infected macrophages [14, 19].

The experimental analysis of selected candidate genes suggested biosynthesis of unsaturated fatty acids as a potential maintenance requirement for *Salmonella* persistency. To more comprehensively evaluate potential targets that are required for this maintenance function, we used an extensively validated genome-scale

computational model of *Salmonella* in vivo metabolism ([31] and manuscript in preparation). Flux-Balance Analysis [32] of the genome-scale metabolic model predicted expected essential genes in the fatty acid biosynthesis pathway (*accA, accB, accC, accD, acpP, fabA, fabD, fabG, fabI*) but no other pathways reflecting redundancy in providing required precursors such as malonyl-CoA, NADPH, and NADH.

Discussion

Chronic infections represent a major health problem. Eradication often requires long-term treatment that causes compliance problems, facilitates resistance development, and often fails to prevent relapse. Many chronic infections are likely to be caused by persistent pathogens in a dormant state with minimal cellular and metabolic activities. In fact, it remains unclear if such dormancy has any basal maintenance requirements that could be targeted for therapy. Various in vitro and in vivo models have been established to determine requirements of persistent pathogens. However, results depend on the particular model and it remains unclear how well these models mimic relevant conditions during chronic infections. It is likely that even within one infected host tissue, various microenvironments exist that might induce distinct forms of persistency [33, 34].

Here we established a simple *Salmonella* mouse infection model in which *Salmonella* with dual metabolic and virulence defects persisted at constant tissue loads. In spleen, a substantial fraction of such *Salmonella* reached a non-proliferating state with minimal cell wall turnover within one day post infection. The fact that a *Salmonella* subpopulation even survived chronic treatment with a fluoroquinolone, the most potent, but still only partially effective antibiotic to eradicate persistent salmonellosis [15] and non-growing bacteria in general [2, 18], indicated that this model represented a stringent test for identifying targets that might be useful in clinically relevant settings. On the other hand, the emergence of small-colony variants of a *fabB* mutant suggested that in this model persistent *Salmonella* still had some metabolic activities that could be diminished to relax residual maintenance requirements.

In striking contrast to *Salmonella* properties in spleen, *Salmonella* in liver remained sensitive to fluoroquinolone treatment and required continues de novo cell wall synthesis. These data suggested that despite purine auxotrophy and inactivity of the SPI-2 type III secretion system, liver microenvironments permitted residual *Salmonella* proliferation in this tissue. Liver colonization was therefore less suitable as readout for *Salmonella* maintenance requirements during persistency. On the other hand, distinct *Salmonella* microenvironments in this organ provided complementary information for target evaluation. As an example, liver seemed to provide conditions that enable at least partial survival of *Salmonella* mutants defective for riboflavin biosynthesis and repair of iron-sulfur clusters, in contrast to conditions in spleen. Antimicrobial chemotherapy should eradicate *Salmonella* from all host organs including liver suggesting that the corresponding targets might be unsuitable.

The *Salmonella* metabolic network contains more than 1200 different enzymes that could all represent potential antimicrobial targets. However, only a very small number of these enzymes are sufficiently important for *Salmonella* physiology to qualify as potentially suitable targets to control acute infections [20]. This problem was even much more severe for persistent infections where we found only a single essential *Salmonella* enzyme indicating strikingly relaxed *Salmonella* requirements for survival as compared to growth in host tissues. Activities that were absolutely essential during acute infection, but dispensable during persistency, include cell wall synthesis, ubiquinone-dependent aerobic respiration, proton motive force-dependent ATP synthesis, translational accuracy, DNA repair, and thioredoxin-mediated redox balance. Dispensability of PMF-driven ATP synthesis highlighted the stringency of our model but

might also reflect differences between *Salmonella* and other pathogens. In addition, activities that play important roles in other persistency models including fatty acid degradation and polyphosphate storage were also dispensable in our model. These data indicate that *Salmonella purA ssaGH* had minimal requirements for extended in vivo survival.

We found only a single defect, inactive biosynthesis of unsaturated fatty acids and cyclopropane fatty acids that resulted in clearance to non-detectable tissue loads within a few days. Interestingly, clearance kinetics for defective mutants were faster compared to the best current antimicrobial drug enrofloxacin for chronic salmonellosis suggesting that the corresponding targets could potentially help to improve treatment of such disease. Unsaturated fatty acids and their derivatives cyclopropane fatty acids together comprise about one-half the *Salmonella* fatty acid content [35]. Essentiality of de novo synthesis could suggest continuous internal turnover, damage, and or loss to the environment. Damage/loss of membranes has previously been proposed as a potential strategy to control persisters [36]. Reactive oxygen species can readily damage mammalian polyunsaturated fatty acids, but bacterial unsaturated fatty acids that usually contain only a single double bond are refractory to oxidative damage [37]. Alternatively, membranes could also be lost by shedding outer membrane vesicles [38]. On the other hand, continuous synthesis of another outer membrane component, lipopolysaccharide may not be needed for *Salmonella* persistence based on the slow clearance of *Salmonella purA ssaGH gutQ yrbH* from infected spleen (Fig. 3). Further studies are needed to clarify the function of de novo fatty acid synthesis during *Salmonella* persistency. It is also important to note that host fatty acids (both saturated

and unsaturated) might be sufficiently available in other infectious disease models, especially in case of extracellular pathogens [39].

In conclusion, we established a practical, highly stringent in vivo persistency model. Data obtained with this model revealed an extremely limited target space for antimicrobial chemotherapy of persistent *Salmonella* infections. On the other hand, at least some *Salmonella* metabolic activities might be crucial for persistency and the model could help to identify additional requirements in subsequent studies.

Materials and Methods

Bacterial genetics

We used strain *Salmonella enterica* serovar typhimurium SL1344 *hisG xyl* [40] as parental wild type strain. *Salmonella* mutants were constructed by lamda red-recombinase mediated allelic replacement [41] followed by general transduction using phage P22 *int* [42]. Resistance cassettes were flanked with FRT sites for removal using FLP recombinase [41]. All strains were cultivated at 37°C in Lennox LB medium containing 90 µg/ml streptomycin and 50µg/ml kanamycin, 20 µg/ml chloramphenicol, and/or 100 µg/ml ampicillin, as appropriate. Auxotrophs were supplemented with 40 µg/ml riboflavin (*ribB*), 0.1% oleate (*fabB*), 50 µg/ml diamino pimelic acid (*asd*), 15µM D-arabinose-5-phosphate / 10 µM glucose-6-phosphate (*gutQ yrbH*). Agar plates containing oleate were always freshly prepared and maintained at 37ºC to keep oleate homogeneously dispersed.

Mouse infections

All animals were handled in strict accordance with good animal practice and all animal work was approved by local animal care and use committee (license 2239, Kantonales Veterinäramt BS). Eight to 12 weeks old female BALB/c mice were infected intravenously with 10^6 CFU *Salmonella* from late exponential LB cultures. For some experiments, we administered enrofloxacin (2 mg/ml) in the drinking water beginning two days post infection [19], or thiolactomycin (two doses of 2 mg per mouse). For

competitive infections, mutant *Salmonella* carrying different antibiotic resistance cassettes were mixed before administration. The actual bacterial dose was confirmed by plating. At various time intervals post infection, mice were sacrificed, spleen and liver collected aseptically in 1ml of 0.1% Triton Tx-100, and number of viable bacteria per organ was determined by plating tissue homogenates on appropriate selective media. Competitive indices (CI=output ratio/input ratio) were calculated based on plate counts for inoculum and tissue homogenates collected at seven days post infection.

In silico modeling

To predict additional targets, we used a genome-scale computational *Salmonella in vivo* metabolism model STMv1.1, an updated version of the consensus genome-scale metabolism reconstruction STMv1 [31] (manuscript in preparation). We used production of unsaturated fatty acids as objective function and determined all genes that were predicted to be essential for this function with Flux-Balance Analysis [32] using MatLab and the COBRA toolbox [43].

Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft (European Research Training Group 1273) and Schweizerischer Nationalfonds (31003A-121834).

References

1. Lewis K: **Persister cells**. *Annu Rev Microbiol* 2010, **64**:357-372.

2. Levin BR, Rozen DE: **Non-inherited antibiotic resistance**. *Nat Rev Microbiol* 2006, **4**(7):556-562.

3. Bryk R, Gold B, Venugopal A, Singh J, Samy R, Pupek K, Cao H, Popescu C, Gurney M, Hotha S *et al*: **Selective killing of nonreplicating mycobacteria**. *Cell Host Microbe* 2008, **3**(3):137-145.

4. Fauvart M, De Groote VN, Michiels J: **Role of persister cells in chronic infections: clinical relevance and perspectives on anti-persister therapies**. *J Med Microbiol*, **60**(Pt 6):699-709.

5. Baek SH, Li AH, Sassetti CM: **Metabolic regulation of mycobacterial growth and antibiotic sensitivity**. *PLoS Biol*, **9**(5):e1001065.

6. John AK, Schmaler M, Khanna N, Landmann R: **Reversible daptomycin tolerance of adherent staphylococci in an implant infection model**. *Antimicrob Agents Chemother*, **55**(7):3510-3516.

7. Dhar N, McKinney JD: **Mycobacterium tuberculosis persistence mutants identified by screening in isoniazid-treated mice**. *Proc Natl Acad Sci U S A*, **107**(27):12275-12280.

8. Munoz-Elias EJ, Timm J, Botha T, Chan WT, Gomez JE, McKinney JD: **Replication dynamics of Mycobacterium tuberculosis in chronically infected mice**. *Infect Immun* 2005, **73**(1):546-551.

9. Rao SP, Alonso S, Rand L, Dick T, Pethe K: **The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis**. *Proc Natl Acad Sci U S A* 2008, **105**(33):11945-11950.

10. Diacon AH, Pym A, Grobusch M, Patientia R, Rustomjee R, Page-Shipp L, Pistorius C, Krause R, Bogoshi M, Churchyard G *et al*: **The diarylquinoline TMC207 for multidrug-resistant tuberculosis**. *N Engl J Med* 2009, **360**(23):2397-2405.

11. Tischler AD, McKinney JD: **Contrasting persistence strategies in Salmonella and Mycobacterium**. *Curr Opin Microbiol*, **13**(1):93-99.

12. Monack DM: **Salmonella persistence and transmission strategies**. *Curr Opin Microbiol* 2011, **2011**:30.

13. Nath G, Singh YK, Maurya P, Gulati AK, Srivastava RC, Tripathi SK: **Does Salmonella Typhi primarily reside in the liver of chronic typhoid carriers?** *J Infect Dev Ctries*, **4**(4):259-261.

14. Monack DM, Bouley DM, Falkow S: **Salmonella typhimurium Persists within Macrophages in the Mesenteric Lymph Nodes of Chronically Infected Nramp1+/+ Mice and Can Be Reactivated by IFN{gamma} Neutralization**. *JExpMed* 2004, **199**(2):231-241.

15. Bhan MK, Bahl R, Bhatnagar S: **Typhoid and paratyphoid fever**. *Lancet* 2005, **366**(9487):749-762.

16. McFarland WC, Stocker BA: **Effect of different purine auxotrophic mutations on mouse-virulence of a Vi-positive strain of Salmonella dublin and of two strains of Salmonella typhimurium**. *MicrobPathog* 1987, **3**(2):129-141.

17. Helaine S, Thompson JA, Watson KG, Liu M, Boyle C, Holden DW: **Dynamics of intracellular bacterial replication at the single cell level**. *Proc Natl Acad Sci U S A* 2010, **107**(8):3746-3751.

18. Eng RH, Padberg FT, Smith SM, Tan EN, Cherubin CE: **Bactericidal effects of antibiotics on slowly growing and nongrowing bacteria**. *Antimicrob Agents Chemother* 1991, **35**(9):1824-1828.

19. Griffin AJ, Li LX, Voedisch S, Pabst O, McSorley SJ: **Dissemination of persistent intestinal bacteria via the mesenteric lymph nodes causes typhoid relapse**. *Infect Immun* 2011, **79**(4):1479-1488.

20. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, Mann M, Bumann D: **Robust Salmonella metabolism limits possibilities for new antimicrobials**. *Nature* 2006, **440**(7082):303-307.

21. Turner AK, Barber LZ, Wigley P, Muhammad S, Jones MA, Lovell MA, Hulme S, Barrow PA: **Contribution of proton-translocating proteins to the virulence of Salmonella enterica serovars Typhimurium, Gallinarum, and Dublin in chickens and mice**. *InfectImmun* 2003, **71**(6):3392-3401.

22. Buchmeier NA, Lipps CJ, So MY, Heffron F: **Recombination-deficient mutants of Salmonella typhimurium are avirulent and sensitive to the oxidative burst of macrophages**. *MolMicrobiol* 1993, **7**(6):933-936.

23. Bjur E, Eriksson-Ygberg S, Aslund F, Rhen M: **Thioredoxin 1 promotes intracellular replication and virulence of Salmonella enterica serovar Typhimurium**. *Infect Immun* 2006, **74**(9):5140-5151.

24. Clements MO, Eriksson S, Thompson A, Lucchini S, Hinton JC, Normark S, Rhen M: **Polynucleotide phosphorylase is a global regulator of virulence and persistency in Salmonella enterica**. *ProcNatlAcadSciUSA* 2002, **99**(13):8784-8789.

25. Fang FC, Libby SJ, Castor ME, Fung AM: **Isocitrate lyase (AceA) is required for Salmonella persistence but not for acute lethal infection in mice**. *InfectImmun* 2005, **73**(4):2547-2549.

26. Grant AJ, Restif O, McKinley TJ, Sheppard M, Maskell DJ, Mastroeni P: **Modelling within-host spatiotemporal dynamics of invasive bacterial disease**. *PLoS Biol* 2008, **6**(4):e74.

27. Meredith TC, Woodard RW: **Identification of GutQ from Escherichia coli as a D-arabinose 5-phosphate isomerase**. *J Bacteriol* 2005, **187**(20):6936-6942.

28. Djaman O, Outten FW, Imlay JA: **Repair of oxidized iron-sulfur clusters in Escherichia coli**. *J Biol Chem* 2004, **279**(43):44590-44599.

29. Machutta CA, Bommineni GR, Luckner SR, Kapilashrami K, Ruzsicska B, Simmerling C, Kisker C, Tonge PJ: **Slow onset inhibition of bacterial beta-ketoacyl-acyl carrier protein synthases by thiolactomycin**. *J Biol Chem* 2009, **285**(9):6161-6169

.

30. Miyakawa S, Suzuki K, Noto T, Harada Y, Okazaki H: **Thiolactomycin, a new antibiotic. IV. Biological properties and chemotherapeutic activity in mice**. *J Antibiot (Tokyo)* 1982, **35**(4):411-419.

31. Thiele I, Hyduke DR, Steeb B, Fankam G, Allen DK, Bazzani S, Charusanti P, Chen FC, Fleming RM, Hsiung CA *et al*: **A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2**. *BMC Syst Biol* 2011, **5**(8):8.

32. Oberhardt MA, Palsson BO, Papin JA: **Applications of genome-scale metabolic reconstructions**. *Mol Syst Biol* 2009, **5**(320):320.

33. Gomez JE, McKinney JD: **M. tuberculosis persistence, latency, and drug tolerance**. *Tuberculosis (Edinb)* 2004, **84**(1-2):29-44.

34. Stewart GR, Robertson BD, Young DB: **Tuberculosis: a problem with persistence**. *Nat Rev Microbiol* 2003, **1**(2):97-105.

35. Dubois-Brissonnet F, Naitali M, Mafu AA, Briandet R: **Induction of fatty acid composition modifications and tolerance to biocides in Salmonella enterica serovar Typhimurium by plant-derived terpenes**. *Appl Environ Microbiol*, **77**(3):906-910.

36. Hurdle JG, O'Neill AJ, Chopra I, Lee RE: **Targeting bacterial membrane function: an underexploited mechanism for treating persistent infections**. *Nat Rev Microbiol* 2011, **9**(1):62-75.

37. Imlay JA: **Pathways of oxidative damage**. *Annu Rev Microbiol* 2003, **57**:395-418.

38. Kulp A, Kuehn MJ: **Biological functions and biogenesis of secreted bacterial outer membrane vesicles**. *Annu Rev Microbiol*, **64**:163-184.

39. Brinster S, Lamberet G, Staels B, Trieu-Cuot P, Gruss A, Poyart C: **Type II fatty acid synthesis is not a suitable antibiotic target for Gram-positive pathogens**. *Nature* 2009, **458**(7234):83-86.

40. Hoiseth SK, Stocker BA: **Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines**. *Nature* 1981, **291**(5812):238-239.

41. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products**. *ProcNatlAcadSciUSA* 2000, **97**(12):6640-6645.

42. Thierauf A, Perez G, Maloy AS: **Generalized transduction**. *Methods Mol Biol* 2009, **501**:267-286.

43. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox**. *Nat Protoc* 2007, **2**(3):727-738.

Figures:

Figure 1

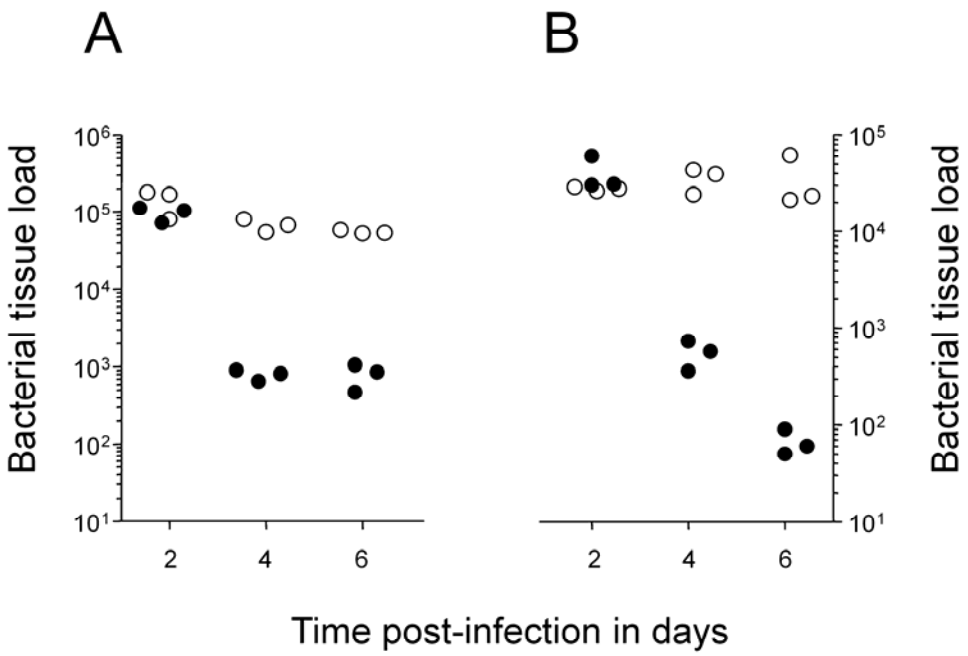


Fig. 1: Colonization kinetics of *Salmonella enterica* serovar Typhimurium *purA ssaGH* in mice. Data are shown for spleen **(A)** and liver **(B)** of individual untreated mice (open circles) and mice that were treated from day two post infection with enrofloxacin (filled circles).

Figure 2

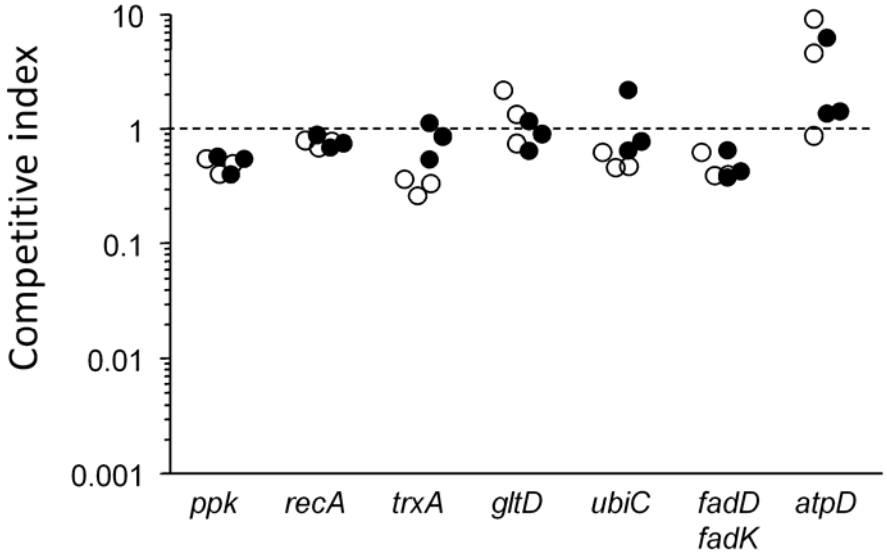


Fig. 2: Competitive indices (CI) of various *Salmonella* mutants vs. the parental *Salmonella purA ssaGH* strain in infected spleen. Data are shown for individual mice at day seven post infection. A competitive index of 1 indicates equal colonization capabilities.

Figure 3

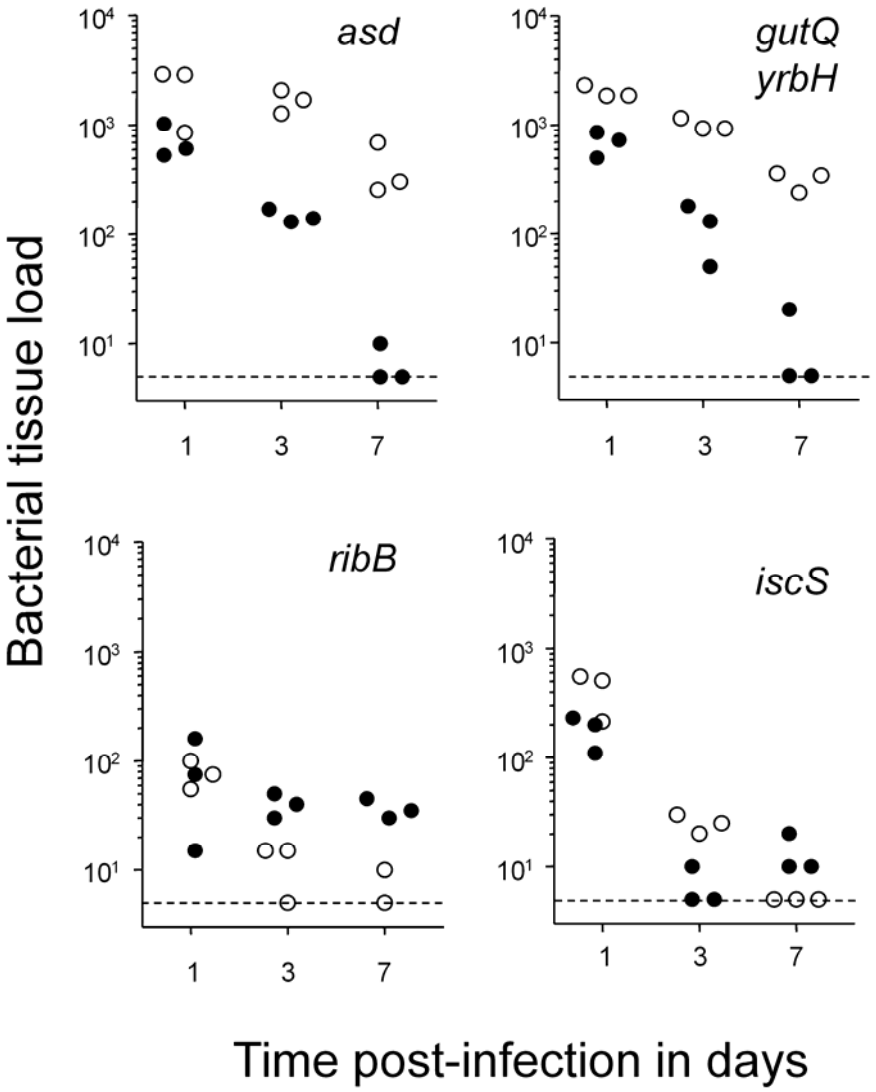


Fig. 3: Colonization kinetics of four compromised mutants in spleen (open circles) and liver (filled circles).

Figure 4

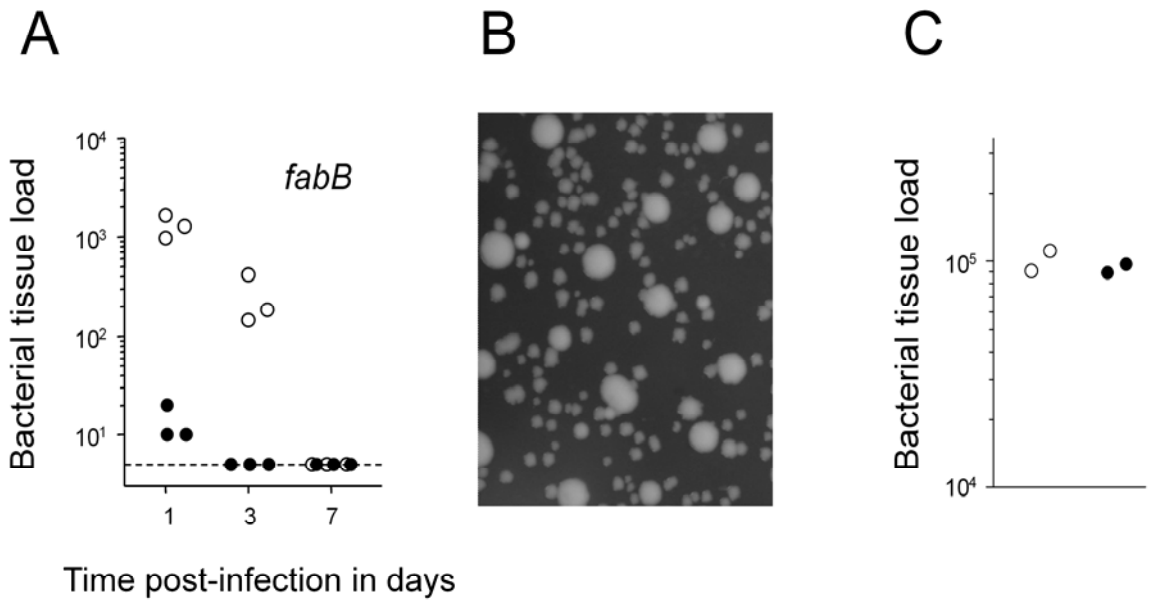


Fig. 4: Clearance of *Salmonella purA ssaGH fabB* from infected mice. A) Colonization kinetics in spleen (open circles) and liver (filled circles). Similar results were obtained in three independent experiments. B) Heterogeneity of colony size on agar plates. Similar data were obtained for two independent in vitro cultures and five independent ex vivo cultures. C) Failure of thiolactomycin treatment of mice infected with Salmonella *purA ssaGH* (open circles, vehicle treated mice; filled circles, thiolactomycin-treated mice).

7.2 List of abbreviations

BiGG database, Biochemically, Genetically and Genomically structured database; **CFU**, Colony forming units; ***dam***, DNA adenine methyltransferase; ***E. coli***, *Escherichia coli*; **FBA**, Flux balance analysis; **InDel**, Insertion/deletion; **LB**, Lysogeny broth; **M cells**, Microfold cells; **MR**, Metabolic reconstruction; **NTS**, Non-typhoidal *Salmonella*; **Rif^{r}**, Rifampicin resistance; ***S.* Enteritis**, *Salmonella enterica* subspecies *enterica* serovar Enteritis; ***S.* Typhi**, *Salmonella enterica* subspecies *enterica* serovar Typhi; ***S.* Typhimurium**, *Salmonella enterica* subspecies *enterica* serovar Typhimurium; **SBML**, Systems biology markup language; **SCV**, *Salmonella* containing vacuole; **SNP**, Single nucleotide polymorphism; **SPI1**, *Salmonella* pathogenicity island 1; **SPI2**, *Salmonella* pathogenicity island 2; **T3SS**, Type III secretion system.