Published in final edited form as:
Curr Opin Biotechnol. 2009 April ; 20(2): 149–157. doi:10.1016/j.copbio.2009.02.002.

Novel genetic tools for studying food borne *Salmonella*

Helene L. Andrews-Polymenis[1], **Carlos A Santiviago**[3], and **Michael McClelland**[2]

[1] Texas A&M University System Health Science Center, College of Medicine, 407 Joe H. Reynolds Medical Building, College Station, TX 77843-1114, handrews@medicine.tamhsc.edu [2] Sidney Kimmel Cancer Center, 10905 Road to the Cure, San Diego CA 92121, mmcclelland@skcc.org [3] Programa de Microbiologia y Micologia, Instituto de Ciencias Biome´dicas (ICBM), Facultad de Medicina, Universidad de Chile, Av. Independencia 1027, Santiago, Chile

Summary of Recent Advances

Non-typhoidal *Salmonellae* are highly prevalent food borne pathogens. High-throughput sequencing of *Salmonella* genomes is expanding our knowledge of the evolution of serovars and epidemic isolates. Genome sequences have also allowed the creation of complete microarrays. Microarrays have improved the throughput of *In vivo* expression technology (IVET) used to uncover promoters active during infection. In another method, signature tagged mutagenesis (STM), pools of mutants are subjected to selection. Changes in the population are monitored on a microarray, revealing genes under selection. Complete genome sequences permit the construction of pools of targeted in-frame deletions that have improved STM by minimizing the number of clones and the polarity of each mutant. Together, genome sequences and the continuing development of new tools for functional genomics will drive a revolution in the understanding of *Salmonellae* in many different niches that are critical for food safety.

Introduction

Non-typhoidal *Salmonella*e are responsible for an estimated 1.4 million cases of gastrointestinal disease with 500 associated deaths in the United States, at a cost of $2 billion [1]. The number of cases worldwide probably exceeds one hundred million each year. Infection generally occurs after ingestion of contaminated food or water, and usually leads to a self-limiting enterocolitis. The disease is characterized by diarrhea, abdominal cramps, nausea, fever, vomiting and headache lasting 7 to 10 days, followed by a longer period of sub-clinical fecal shedding. Infants, the elderly, and immunocompromised individuals are at risk for serious systemic complications and death as a result of infection.

Contaminated foods, including beef, pork, poultry and egg products are frequent vectors responsible for transmission of these organisms to humans. Livestock can harbor *Salmonellae* sub-clinically resulting in carcass contamination at slaughter and in the laying of contaminated eggs. In recent years, as the traditional routes of infection are better controlled, large outbreaks of non-typhoidal *Salmonella* infection in the United States have been attributed to fruits, vegetables and processed foods including jalapeno peppers, cantaloupe, Malto-meal™ cereal, and peanut butter (http://www.cdc.gov/salmonella/).

Serology based on surface antigens is the standard method of classification of *Salmonella*. The host-range and disease can differ considerably between serovars, making such classification

Correspondence to: Helene L. Andrews-Polymenis.

important. Throughout the world, the most prevalent non-typhoidal serovars isolated from human sources are serovars Typhimurium and Enteritidis and these two serovars comprise nearly 40% of isolations from human sources in the United States [2]. These serovars can be harbored sub-clinically in livestock for prolonged periods of time and are thus very difficult to eradicate in the absence of a detailed knowledge of the biology of the organism in this niche.

The bacterial factors necessary for *Salmonellae* to persist sub-clinically in the gastrointestinal tract of livestock and to survive and grow in other reservoirs such as crops and processed foods is only beginning to be elucidated. This knowledge will allow the development of new strategies and the identification of points in the production chain where producers can intervene to improve the safety foods. We review the current status as well as the uses of complete genome sequence information for *Salmonellae*, and enhancements of genetic techniques that may rapidly increase our knowledge of the biology of this organism in these important food safety niches.

Complete Genome Sequencing of *Salmonellae*

Currently, the complete genome sequences of one or more representatives of six serovars have been determined, annotated, and published [3–8], and additional genomes in seven other serovars are complete (Table 1). Cheaper and faster sequencing technologies, such as 454 and Illumina Solexa [9], are now being applied to extensive sequence comparison of non-typhoidal *Salmonellae* (Table 1). Use of these tools provides a new window into genetic diversity within and between *Salmonella* serovars (and within and between epidemic and non-epidemic isolates) at a level that has not previously been possible. These sequences also provide a scaffold for functional genomic studies of *Salmonella* in particular environments, including livestock models and other sources of food borne infection.

Complete genome sequences have allowed the development of open reading frame (ORF) microarrays [10,11] and complete tiling arrays for *Salmonellae* [12]. Comparative Genomic Hybridization has been performed using microarrays to characterize the gene content of non-sequenced clinical and epidemic isolates that are commonly implicated in human food-borne outbreaks [13–16]. One conclusion from such work is that while most serovars consist of strains that are very similar to each other and differ in gene content from other serovars, there are exceptions to this rule. Some individual serovars consist of strains that differ quite considerably from each other in gene content, whereas, some serovars are very closely related and have almost identical gene content [17]. Thus, DNA-based classification of strains is likely to be important to further refine the host range and disease symptoms associated with particular genome variants (called “genovars”).

Several classification methods have been developed to complement serology. Examples of tools for classification at the serovar and genovar level include multilocus enzyme electrophoresis (MLEE [18]), and DNA-based methods including multilocus sequence typing (MLST [19,20], and multiplex PCR targeted to genes specific to subsets of serovars or genovars [21,22]. The DNA-based methods pulsed-field gel electrophoresis (PFGE) [23–25] and variable number of tandem repeat analysis (VNTR, or Multiple Loci VNTR Analysis (MLVA)) [26,27] detect more rapid genomic changes making these techniques useful for epidemiology. Genome-wide single nucleotide polymorphism analysis is just beginning to be used for genomic typing [28,29], and will likely be used heavily in the future.

Genetic Approaches to Find Genes Needed During Colonization of Foods

A better understanding of the *Salmonella* genes that are required for the colonization and persistence in foods will allow us focus future safety and HACCP programs. Complete genome sequencing and annotation can suggest a likely function for some genes. However, other

methods must be used to determine the function or molecular role of most genes and the particular conditions where each gene is important. Strategies currently in use for the identification of *Salmonella* genes involved in colonization of specialized niches such as livestock and food products generally involve two basic approaches: determination of genes expressed under particular conditions and forward genetic analysis of *Salmonella* mutants.

Gene expression

Determining genes that are expressed in a particular condition or environment has been used as a first step to define groups of candidate genes that are necessary for survival and growth. Microarray technology has been used to determine gene expression in particular environments, some of which are relevant to food safety [30–33], but thus far this technique has not been directly used to define *Salmonella* genes expressed in or on foods. Expression analysis using RNA has disadvantages, particularly the difficulty of obtaining *Salmonella* RNA from highly complex environments, including livestock, animal carcasses, live shellfish and produce. However, methods to capture such RNA for analysis have been developed, including strategies that can enrich RNA from very few bacteria, including purification from inside eukaryotic cells by capturing PCR products of RNAs using the bacterial genome [34].

Another strategy that sidesteps the issue of RNA abundance is *in vivo* expression technology (IVET), which identifies active promoters in a specific environment (for a comprehensive review see [35]). IVET is a promoter trap strategy employing a library of random genomic fragments ligated to a promoterless reporter gene. This library is used to identify promoters that are differentially active *in vivo* but not *in vitro* by assaying the transcriptional activity of promoter-reporter fusions. The bacteria carrying these promoter clones can then be expanded *in vitro* for characterization of the promoter by sequencing.

IVET was adapted for the *in vivo* identification of *S. enterica* genes expressed in mice [36]. Reporter constructs used in IVET include promoterless genes essential for growth (*purA*) [36], antibiotic resistance cassettes [37], recombinase-based systems (RIVET) [38], and promoterless GFP [39,40]. IVET and more recent modified protocols including differential fluorescence induction (DFI) and recombinase-based IVET have been used extensively for identification of *Salmonella in vivo* induced genes in mice and macrophages [41–44]. One of the few applications of this method to a food safety issue is by Huang et al., who recently used a recombinase-based *in vivo* expression system coupled with a PCR- based method to rapidly identify activated promoters, to identify genes expressed during infection of swine [45]. 31 genes were identified in this screen that were expressed during colonization of the porcine intestine and/or tonsil including several known adhesins and colonization factors (*bcfA*, *hscA*, *rffG*), that are likely necessary for attachment to and colonization of the epithelial surfaces in the porcine intestine. Furthermore *yciR*, a diguanylate cyclase phosphodiesterase motif containing protein (GDDEF-EAL) previously shown to be important for the ability of *Salmonellae* to form biofilms [46], was also induced in the porcine intestine [45]. This finding suggests that biofilm formation may be involved in colonization of the swine intestine. Thus, control of *Salmonella* biofilm formation may be a future area for the development of novel approaches to increase food safety. Recently, the ability to make discoveries using IVET methods has been accelerated by the hybridization of mixtures of hundreds or thousands of GFP-expressing clones to a tiling microarray that can identify each individual region represented in the mixture [44]. Continuing improvements may encourage greater use of these methods in pathogenic bacteria that contaminate food.

Finally, chromatin immunoprecipitation (ChIP) is a method for identifying targets of DNA-binding regulatory proteins. In this technique, proteins are cross-linked to DNA in live cells, and antibody specific for a regulatory protein of interest is used to isolate DNA fragments binding that protein. Bound DNA is amplified and characterized using DNA tiling arrays.

Tiling arrays of the *Salmonella* genome have allowed a characterization of the binding patterns of particular regulatory proteins [12]. ChIP and other such approaches will be critical in generating a systems biology description of *Salmonella* as a pathogen.

Genes required in particular environments

Even though expression analysis and IVET identify genes expressed in a particular environment, these techniques cannot define genes that are required to colonize a given ecological niche. Furthermore, required genes may be only transiently expressed at low levels, and thus may be missed by some RNA expression and IVET-based strategies. A more direct method for finding the subset of required genes is to use forward genetic screening.

Signature Tagged Mutagenesis (STM) is a negative selection strategy developed to identify virulence factors of *Salmonella enterica* serovar Typhimurium in mice [47]. In the original version of STM, a collection of mutants generated with uniquely marked transposons was pooled and passed through a selective condition. The unique tags present in the input pool but missing in the output pools identify mutants that are unable to survive in the selective condition of interest. Such mutants are identified by hybridization to arrays of signature tags [47,48]. More recently the internal tag in the transposon has been replaced by using transcripts of the unique genome sequence adjacent each transposon generated from a T7 promoter located inside the transposon [49–51]. STM combines the advantages of transposon mutagenesis with the ability to screen a larger number of mutants using fewer animals, a factor that is critically important when using livestock models that are cumbersome and expensive.

Signature tagged mutagenesis has been used much more extensively than IVET to identify genes in various *Salmonella* serovars necessary for colonization of livestock that are the primary sources of contaminated meat and poultry products consumed by humans. In publications that each use a few hundred random transposon mutants for STM, a number of *Salmonella* candidate genes necessary for colonization of calves, chickens and swine have been observed in the broad host range serovar Typhimurium and the narrow host range isolates Gallinarum, Cholerasuis and Dublin that are much less frequently studied [48,52–55]. For example, the *Salmonella* pathogenicity islands 1 and 2 (SPI-1, SPI-2) are needed to colonize the intestinal epithelium of both cattle and swine [48,52]. These studies are also beginning to outline the genes necessary only in a particular host. For example, the genes of SPI-4 are required for colonization of the bovine intestine, but are not required for colonization in swine [48,52]. Genes of SPI-6, in contrast, are necessary for colonization in swine but not in calves. The genes of the major pathogenicity islands SPI-1 and SPI-2 are required for colonization of both bovine and porcine intestinal epithelium, but are not required for colonization of intestinal contents in poultry [48,52]. Genetic requirements for *Salmonella* growth and survival have also been examined for less complex conditions that are relevant to food safety using signature tagged mutagenesis [56]. Exquisitely sensitive techniques such as STM should also be used to investigate the ecology of *Salmonella* and *E. coli* growth on other foods, including those consumed raw.

Recent Advances in Functional Genomics of *Salmonellae*

One limitation of using random transposon mutagenesis is that very large numbers of mutants are needed to ensure complete coverage of the genome. This factor is a significant disadvantage in circumstances where the population of bacteria experiences random loss, also termed a ‘founder effect’ or bottleneck. For example, when a population of *Salmonellae* passes from the intestine to systemic sites only a small fraction of the bacteria arrive in the new niche. Thus, only a few hundred or a few thousand bacteria can be pooled in circumstances where founder effects lead to random loss of mutants from the population. Furthermore, transposon mutants

can have polar effects on downstream gene expression, and occasionally on upstream gene expression by producing interfering transcripts.

To circumvent these limitations, we have employed complete genome sequence information to generate ordered libraries of targeted non-polar deletion mutants of individual genes using the lambda-red PCR-recombination method of Datsenko & Wanner [57]. We have developed a strategy in which pools of these specific knockout mutants are studied using a negative selection strategy similar to STM (Figure 2) (Santiviago et al. unpublished). We have also constructed defined mutants that delete multiple adjacent genes, thereby reducing the complexity of a pooled mutant library even further (Santiviago et al., unpublished). Similar to other recent STM strategies, the mutants present in the input but selected for or against in the output pools are identified using microarray analysis.

The main advantages of this approach for screening in complex environments such as livestock, is the 10-fold or even 100-fold reduced complexity of the pool of defined mutants relative to the complexity of a random transposon pool needed for equivalent coverage of the genome. Other advantages include the reduction of polar effects, and the existence of targeted clean deletion mutants for confirmation of phenotypes identified by screening. These improvements enhance the ability to efficiently and accurately detect which *Salmonella* genes have a role in a host of diverse environments and allow more comprehensive screening of the *Salmonella* genome in complex environments than has previously been possible. The methods should be adaptable to other food borne bacterial pathogens.

Finally, as the cost of sequencing continues to plummet, and is combined with the ability to multiplex many biological samples in one sequencing run, it is possible that sequencing will replace microarrays for determining the population structure of RNAs [58], for characterizing protein DNA-complexes in chromatin immunoprecipitation [59], and for monitoring mutants.

The Future

Complete genome sequencing of *Salmonellae* is allowing us to better understand their genetic diversity, to develop novel tools, and to improve existing genetic techniques to understand the complex biology of these important food borne pathogens. Approximately half of the genes in *Salmonella* still have no known phenotype in the environment. Frontiers for further study of *Salmonella* for improved food safety using modern genetic tools are likely to include determination of the genes necessary for environments where *Salmonella* must survive outside the host, such as in feces, soil, water, and plants. Understanding how *Salmonella* completes its entire host-to-host life cycle in agriculture may reveal previously unknown vulnerabilities that will be susceptible to novel intervention and allow us to break the chain of transmission.

Acknowledgments

This work was supported in part by NIH grants R01 AI034829, and R01 AI052237, and R21AI057733, DOD grant BC073899, by the Texas A&M University System Health Science Center, and the generous support of Mr. Sidney Kimmel.

References

1. Frenzen PRT, Buzby J. Salmonella cost estimate updated using FoodNet data. Food Review 1999;22:10–15.
2. CDC. Salmonella Annual Summary 2005. Edited by: Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute of Infectious Diseases; 2005.
3. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, et al. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature 2001;413:852–856. [PubMed: 11677609]

4. 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. [PubMed: 11677608]

5. McClelland M, Sanderson KE, Clifton SW, Latreille P, Porwollik S, Sabo A, Meyer R, Bieri T, Ozersky P, McLellan M, et al. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of Salmonella enterica that cause typhoid. Nat Genet. 2004

6. Deng W, Liou SR, Plunkett G 3rd, Mayhew GF, Rose DJ, Burland V, Kodoyianni V, Schwartz DC, Blattner FR. Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. J Bacteriol 2003;185:2330–2337. [PubMed: 12644504]

7. Chiu CH, Tang P, Chu C, Hu S, Bao Q, Yu J, Chou YY, Wang HS, Lee YS. The genome sequence of Salmonella enterica serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. Nucleic Acids Res 2005;33:1690–1698. [PubMed: 15781495]

**8. Thomson NR, Clayton DJ, Windhorst D, Vernikos G, Davidson S, Churcher C, Quail MA, Stevens M, Jones MA, Watson M, et al. Comparative genome analysis of Salmonella Enteritidis PT4 and Salmonella Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. Genome Res 2008;18:1624–1637. [PubMed: 18583645]Comparison of genome sequences determines that chicken-restricted serovar Gallinarum is a recent descendent of the wider host range serovar Enteritidis. This paper provides a basis for directed study of host restriction in *Salmonellae*

*9. 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. [PubMed: 18660809]First high-throughput sequencing to determined the complete genome sequence for large numbers of *Salmonella*: 19 isolates of S. Typhi, representing major nodes in the phylogenetic tree.

10. Porwollik S, Wong RM, McClelland Michael. Evolutionary genomics of *Salmonella*: Gene acquisitions revealed by microarray analysis. Proceedings of the National Academy of Sciences 2002;99:8956–8961.

11. Chan K, Baker S, Kim CC, Detweiler CS, Dougan G, Falkow S. Genomic comparison of Salmonella enterica serovars and Salmonella bongori by use of an S. enterica serovar typhimurium DNA microarray. J Bacteriol 2003;185:553–563. [PubMed: 12511502]

12. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC. Selective Silencing of Foreign DNA with Low GC Content by the H-NS Protein in Salmonella. Science 2006;313:236–238. [PubMed: 16763111]

13. Porwollik S, Santiviago CA, Cheng P, Florea L, McClelland M. Differences in gene content between Salmonella enterica serovar Enteritidis isolates and comparison to closely related serovars Gallinarum and Dublin. J Bacteriol 2005;187:6545–6555. [PubMed: 16159788]

14. Kang MS, Besser TE, Hancock DD, Porwollik S, McClelland M, Call DR. Identification of specific gene sequences conserved in contemporary epidemic strains of Salmonella enterica. Appl Environ Microbiol 2006;72:6938–6947. [PubMed: 16963552]

15. Reen FJ, Boyd EF, Porwollik S, Murphy BP, Gilroy D, Fanning S, McClelland M. Genomic comparisons of Salmonella enterica serovar Dublin, Agona, and Typhimurium strains recently isolated from milk filters and bovine samples from Ireland, using a Salmonella microarray. Appl Environ Microbiol 2005;71:1616–1625. [PubMed: 15746367]

16. Andrews-Polymenis HL, Rabsch W, Porwollik S, McClelland M, Rosetti C, Adams LG, Baumler AJ. Host restriction of Salmonella enterica serotype Typhimurium pigeon isolates does not correlate with loss of discrete genes. J Bacteriol 2004;186:2619–2628. [PubMed: 15090502]

17. Porwollik S, Boyd EF, Choy C, Cheng P, Florea L, Proctor E, McClelland M. Characterization of Salmonella enterica subspecies I genovars by use of microarrays. J Bacteriol 2004;186:5883–5898. [PubMed: 15317794]

18. Boyd EF, Wang F-S, Whittam TS, Selander RK. Molecular genetic relationship of the *Salmonellae*. Appl Environ Microbiol 1996;62:804–808. [PubMed: 8975610]

19. Tankouo-Sandjong B, Sessitsch A, Liebana E, Kornschober C, Allerberger F, Hachler H, Bodrossy L. MLST-v, multilocus sequence typing based on virulence genes, for molecular typing of Salmonella enterica subsp. enterica serovars. Journal of Microbiological Methods 2007;69:23–36. [PubMed: 17208323]

20. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America 1998;95:3140–3145. [PubMed: 9501229]

21. Arrach N, Porwollik S, Cheng P, Cho A, Long F, Choi SH, McClelland M. Salmonella serovar identification using PCR-based detection of gene presence and absence. J Clin Microbiol. 2008

22. Kim S, Frye JG, Hu J, Fedorka-Cray PJ, Gautom R, Boyle DS. Multiplex PCR-based method for identification of common clinical serotypes of Salmonella enterica subsp. enterica. J Clin Microbiol 2006;44:3608–3615. [PubMed: 16943358]

23. Jackson CR, Fedorka-Cray PJ, Barrett JB, Hiott LM, Woodley TA. Prevalence of streptogramin resistance in enterococci from animals: identification of vatD from animal sources in the USA. International Journal of Antimicrobial Agents 2007;30:60–66. [PubMed: 17532190]

**24. Swaminathan B, Barrett TJ, Fields P. Surveillance for human Salmonella infections in the United States. Journal of AOAC International 2006;89:553–559. [PubMed: 16640306]A report from the front line regarding the tools used in food safety to battle *Salmonella*

25. Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB, Swaminathan B, Barrett TJ. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of Escherichia coli O157:H7, Salmonella, and Shigella for PulseNet. Foodborne Pathogens & Disease 2006;3:59–67. [PubMed: 16602980]

26. Cho S, Whittam TS, Boxrud DJ, Bartkus JM, Saeed AM. Allele distribution and genetic diversity of VNTR loci in Salmonella enterica serotype Enteritidis isolates from different sources. BMC Microbiology 2008;8:146. [PubMed: 18793420]

27. Boxrud D, Pederson-Gulrud K, Wotton J, Medus C, Lyszkowicz E, Besser J, Bartkus JM. Comparison of multiple-locus variable-number tandem repeat analysis, pulsed-field gel electrophoresis, and phage typing for subtype analysis of Salmonella enterica serotype Enteritidis. Journal of Clinical Microbiology 2007;45:536–543. [PubMed: 17151203]

28. Octavia S, Lan R. Single-Nucleotide-Polymorphism Typing and Genetic Relationships of Salmonella enterica Serovar Typhi Isolates. J Clin Microbiol 2007;45:3795–3801. [PubMed: 17728466]

29. Baker S, Holt KE, van de Vosse E, Roumagnac P, Whitehead S, King E, Ewels P, Keniry A, Weill FX, Lightfoot D, et al. High-Throughput Genotyping of Salmonella enterica Serovar Typhi Allowing Geographical Assignment of Haplotypes and Pathotypes within an Urban District of Jakarta, Indonesia. J Clin Microbiol 2008;46:1741–1746. [PubMed: 18322069]

30. Wang Q, Mariconda S, Suzuki A, McClelland M, Harshey RM. Uncovering a large set of genes that affect surface motility in Salmonella enterica serovar Typhimurium. Journal of Bacteriology 2006;188:7981–7984. [PubMed: 16980469]

31. Ledeboer NA, Frye JG, McClelland M, Jones BD. Salmonella enterica serovar Typhimurium requires the Lpf, Pef, and Tafi fimbriae for biofilm formation on HEp-2 tissue culture cells and chicken intestinal epithelium. Infect Immun 2006;74:3156–3169. [PubMed: 16714543]

32. Wang Q, Frye JG, McClelland M, Harshey RM. Gene expression patterns during swarming in Salmonella typhimurium: genes specific to surface growth and putative new motility and pathogenicity genes. Mol Microbiol 2004;52:169–187. [PubMed: 15049819]

*33. Frye J, Karlinsey JE, Felise HR, Marzolf B, Dowidar N, McClelland M, Hughes KT. Identification of new flagellar genes of Salmonella enterica serovar Typhimurium. Journal of Bacteriology 2006;188:2233–2243. [PubMed: 16513753]RNA expression data from judiciously selected mutants reveals new genes involved in a well-studied regulatory pathway.

*34. Faucher SP, Porwollik S, Dozois CM, McClelland M, Daigle F. Transcriptome of Salmonella enterica serovar Typhi within macrophages revealed through the selective capture of transcribed sequences. Proceedings of the National Academy of Sciences of the United States of America 2006;103:1906–1911. [PubMed: 16443683]Capture of *Salmonella* RNA by hybridization to the *Salmonella* genome allows characterization of expression from an environment where this organism is a tiny fraction of the mass.

35. Rediers H, Rainey PB, Vanderleyden J, De Mot R. Unraveling the Secret Lives of Bacteria: Use of In Vivo Expression Technology and Differential Fluorescence Induction Promoter Traps a Tools for

Exploring Niche-Specific Gene Expression. Microbiology and Molecular Biology Reviews 2005;69:217–261. [PubMed: 15944455]

36. Mahan MJ, Slauch JM, Mekalanos JJ. Selection of bacterial virulence genes that are specifically induced in host tissues. Science 1993;259:686–688. [PubMed: 8430319]
37. Mahan MJ, Tobias JW, Slauch JM, Hanna PC, Collier RJ, Mekalanos JJ. Antibiotic-based selection for bacterial genes that are specifically induced during infection of a host. Proc Natl Acad Sci U S A 1995;92:669–673. [PubMed: 7846034]
38. Merighi M, Ellermeier CD, Slauch JM, Gunn JS. Resolvase-in vivo expression technology analysis of the Salmonella enterica serovar Typhimurium PhoP and PmrA regulons in BALB/c mice. J Bacteriol 2005;187:7407–7416. [PubMed: 16237024]
39. Valdivia RH, Falkow S. Bacterial genetics by flow cytometry: rapid isolation of Salmonella typhimurium acid-inducible promoters by differential fluorescence induction. Mol Microbiol 1996;22:367–378. [PubMed: 8930920]
40. Bumann D. Examination of Salmonella gene expression in an infected mammalian host using the green fluorescent protein and two-colour flow cytometry. Mol Microbiol 2002;43:1269–1283. [PubMed: 11918812]
41. Stanley TL, Ellermeier CD, Slauch JM. Tissue-specific gene expression identifies a gene in the lysogenic phage Gifsy-1 that affects Salmonella enterica serovar typhimurium survival in Peyer's patches. J Bacteriol 2000;182:4406–4413. [PubMed: 10913072]
42. Valdivia RH, Falkow S. Fluorescence-based isolation of bacterial genes expressed within host cells. Science 1997;277:2007–2011. [PubMed: 9302299]
43. Heithoff DM, Conner CP, Hanna PC, Julio SM, Hentschel U, Mahan MJ. Bacterial infection as assessed by in vivo gene expression. Proc Natl Acad Sci U S A 1997;94:934–939. [PubMed: 9023360]

**44. Arrach N, Zhao M, Porwollik S, Hoffman RM, McClelland M. Salmonella promoters preferentially activated inside tumors. Cancer Res 2008;68:4827–4832. [PubMed: 18559530]First use of a promoter trap-GFP library and an oligonucleotide tiling array: Used to identify *Salmonella* promoters expressed in tissues in mice.

45. Huang Y, Leming CL, Suyemoto M, Altier C. Genome-Wide Screen of Salmonella Genes Expressed during Infection in Pigs Using In Vivo Expression Technology. Appl Environ Microbiol 2007;73:7522–7530. [PubMed: 17921269]
46. Garcia B, Latasa C, Solano C, Garcia-del Portillo F, Gamazo C, Lasa I. Role of the GGDEF protein family in Salmonella cellulose biosynthesis and biofilm formation. Mol Microbiol 2004;54:264–277. [PubMed: 15458421]
47. 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. [PubMed: 7618105]
48. Morgan E, Campbell JD, Rowe SC, Bispham J, Stevens MP, Bowen AJ, Barrow PA, Maskell D, Wallis TS. Identification of host-specific colonization factors of *Salmonella enterica* serovar Typhimurium. Mol Microbiol 2004:53. [PubMed: 15049810]

*49. Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM. Genome-wide screen for salmonella genes required for long-term systemic infection of the mouse. PLoS Pathog 2006;2:e11. [PubMed: 16518469]First study of long-term systemic carriage using transposons mutants carrying the T7 promoter and an ORF microarray,

50. Chan K, Kim CC, Falkow S. Microarray-based detection of Salmonella enterica serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. Infect Immun 2005;73:5438–5449. [PubMed: 16113260]
51. Badarinarayana V, Estep PW 3rd, Shendure J, Edwards J, Tavazoie S, Lam F, Church GM. Selection analyses of insertional mutants using subgenic-resolution arrays. Nat Biotechnol 2001;19:1060–1065. [PubMed: 11689852]
52. Carnell SC, Bowen AJ, Morgan E, Maskell D, Wallis TS, Stevens MP. Role in virulence and protective efficacy in pigs of Salmonella enterica serovar Typhimurium secreted components identified by signature-tagged mutagenesis. Microbiology 2007;153:1940–1952. [PubMed: 17526851]Signature Tagged Mutagenesis screening of a library of 1045 random transposon mutants used to identify

Salmonella candidate genes necessary for intestinal colonization in swine. This study illustrates of the complexity of such work in livestock.

53. Bispham J, Tripathi BN, Watson PR, Wallis TS. Salmonella pathogenicity island 2 influences both systemic salmonellosis and Salmonella-induced enteritis in calves. Infect Immun 2001;69:367–377. [PubMed: 11119526]
54. Shah DH, Lee MJ, Park JH, Lee JH, Eo SK, Kwon JT, Chae JS. Identification of Salmonella gallinarum virulence genes in a chicken infection model using PCR-based signature-tagged mutagenesis. Microbiology 2005;151:3957–3968. [PubMed: 16339940]
55. Ku YW, McDonough SP, Palaniappan RU, Chang CF, Chang YF. Novel attenuated Salmonella enterica serovar Choleraesuis strains as live vaccine candidates generated by signature-tagged mutagenesis. Infection & Immunity 2005;73:8194–8203. [PubMed: 16299315]
56. Bearson SM, Bearson BL, Rasmussen MA. Identification of Salmonella enterica serovar Typhimurium genes important for survival in the swine gastric environment. Appl Environ Microbiol 2006;72:2829–2836. [PubMed: 16597989]
57. 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. [PubMed: 10829079]
58. Hoen PAC, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RHAM, de Menezes RX, Boer JM, van Ommen GJB, den Dunnen JT. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Res 2008;2008:e141. [PubMed: 18927111]
59. Wederell ED, Beilenky M, Cullum R, Thiessen N, Dagpinar M, Delaney A, Varhol R, Zhao YJ, Zeng T, Marra MA, et al. Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. Nucleic Acids Res 2008;36:4549–4564. [PubMed: 18611952]

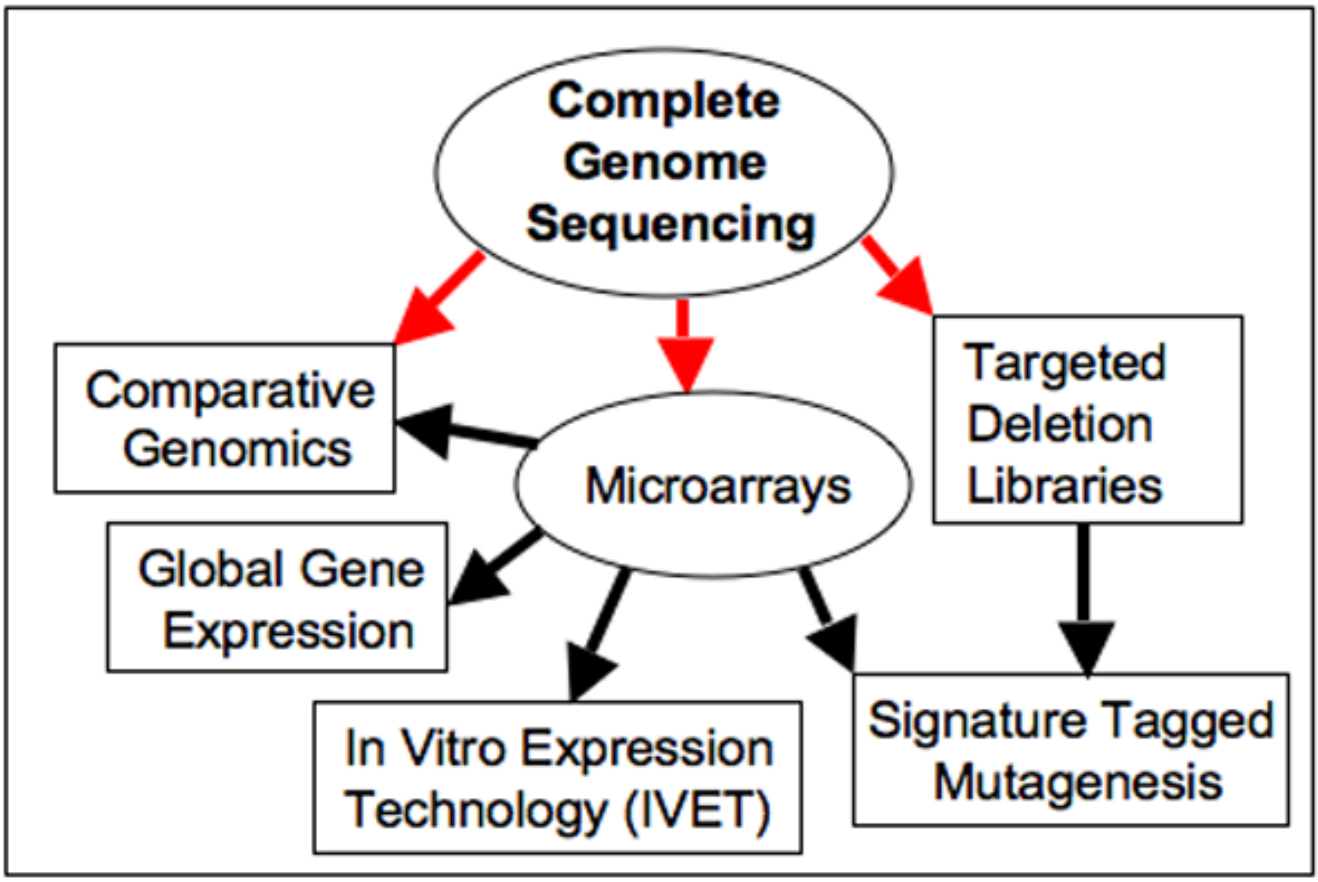

Figure 1. The Influence of Complete Genome Sequencing on *Salmonella* Genetics

Complete genome sequencing has revolutionized comparative genomics of *Salmonellae*, and allowed the development of DNA microarrays and targeted deletion libraries. These tools are accelerating both the accuracy and the coverage obtained in gene expression studies and forward genetic analysis of mutants.

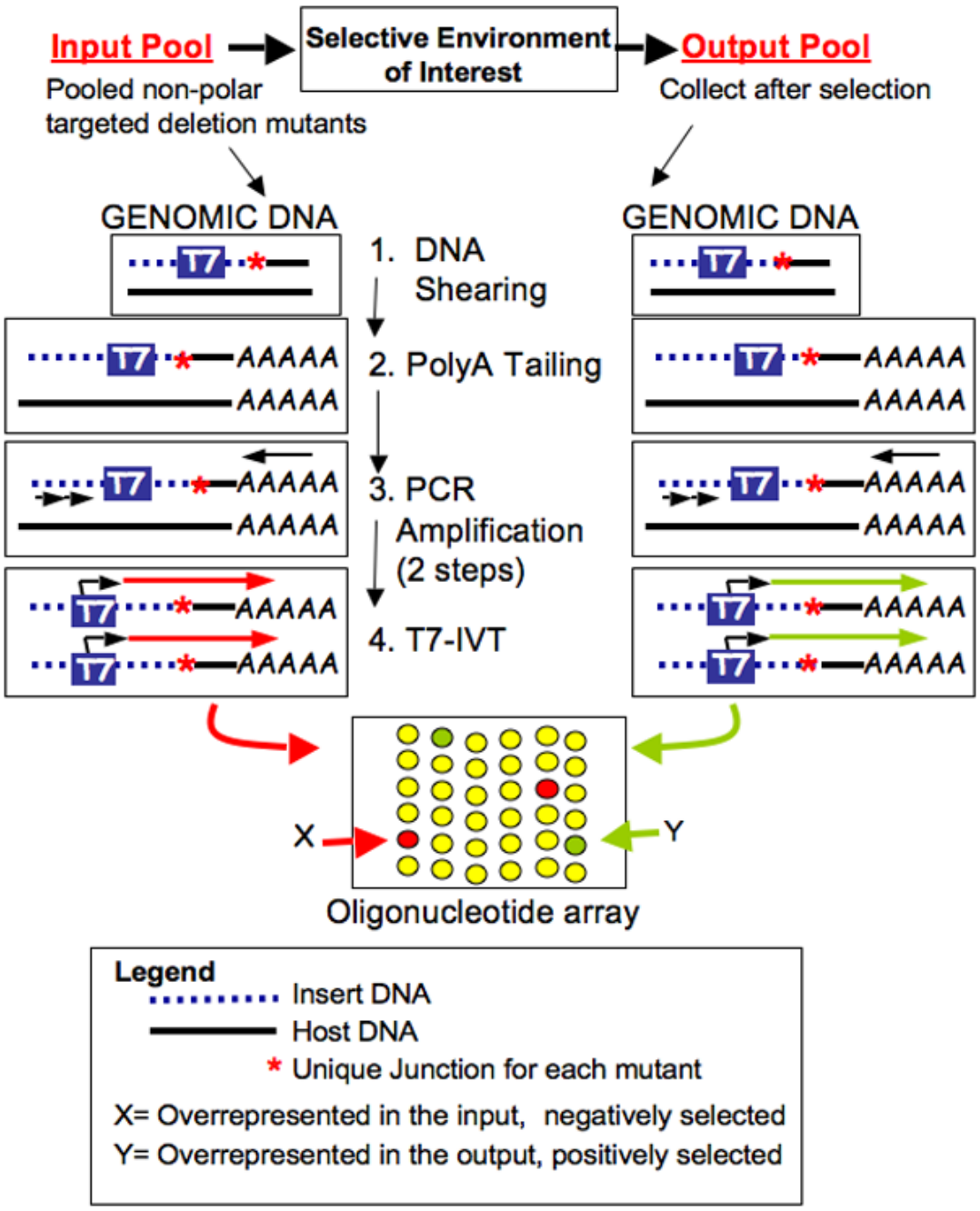

Figure 2. Forward Genetic Method to Identify *Salmonella* Genes Selected in an Environment of Interest

Forward Genetic Method to Identify *Salmonella* Genes Selected in an Environment of Interest. Libraries of targeted deletion mutants are constructed. Mutants are passed through a selective condition as a pool. Genomic DNA from both the input pool and the output pool is sheared, polyadenylated, and nested PCR is used to specifically amplify junction fragments containing the T7 promoter. The resulting amplified product is used for *T7 in vitro* transcription with direct incorporation of fluorescent nucleotides. Mutants selected for or against are determined by comparison of the labeled transcripts in the input pool to the labeled transcripts in the output pool using an oligonucleotide microarray of genomic sequences directly adjacent each mutant.

Table 1

Current Status of Genome Sequencing for *Salmonellae*

Sequences Completed	Method	Sequencing Center	GenBank Accession #	Citation
* Arizonae 62:z4,z23:--	ABI	WU	CP000880.1	
Agona SL483	ABI	JCVI	CP001138.1	
Choleraesuis SC-B67	ABI	Taiwan	AE017220.1	[7]
Dublin CT_02021853	ABI	TIGR	CP001144.1	
Enteritidis P125109	ABI	Sanger	AM933172.1	[8]
Gallinarum 287/91	ABI	Sanger	AM933173.1	[8]
Heidelberg SL476	ABI	JCVI	CP001120.1	
Newport SL254	ABI	TIGR/JCVI	CP001113.1	
Paratyphi A ATCC 9150	ABI	WU	CP000026.1	[5]
Paratyphi B SPB7 SGSC4150	ABI	WU	CP000886.1	
Schwarzengrund CVM19633	ABI	TIGR	CP001127.1	
Typhi E98-0664/Kenya 1998i	454	Sanger	NZ_CAAU00000000	[9]
Typhi 150(98)S/Vietnam 2004	Illumina GS	Sanger		[9]
Typhi 404ty/Indonesia 1983	454/Illumina GS	Sanger	NZ_CAAQ00000000	[9]
Typhi 8(04)N/Vietnam 2004	Illumina GS	Sanger		[9]
Typhi AG3/Vietnam 1998	454/Illumina GS	Sanger	NZ_CAAY00000000	[9]
Typhi CT18/Vietnam 1993	ABI Illumina GS	Sanger	AL513382.1	[4] [9]
Typhi E00-7866/Morocco 2000	454	Sanger	NZ_CAAR00000000	[9]
Typhi E01-6750/Senegal 2001	454	Sanger	NZ_CAAS00000000	[9]
Typhi E02-1180/India 2002	454	Sanger	NZ_CAAT00000000	[9]
Typhi E02-2759/India 2002	Illumina GS	Sanger		[9]
Typhi E03-4983/Indonesia 2003	Illumina GS	Sanger		[9]
Typhi E03-9804/Nepal 2003	Illumina GS	Sanger		[9]
Typhi E98-2068/Bangladesh 1998	454	Sanger	NZ_CAAV00000000	[9]
Typhi E98-3139/Mexico 1998	454/Illumina GS	Sanger	NZ_CAAZ00000000	[9]
Typhi ISP-03-07467/Morocco 2003	Illumina GS	Sanger		[9]
Typhi ISP-04-06979/Africa 2004	Illumina GS	Sanger		[9]
Typhi J185SM/Indonesia 1985	454	Sanger	NZ_CAAW00000000	[9]
Typhi M223/Unknown 1939	454	Sanger	NZ_CAAX00000000	[9]
Typhi Ty2 Russia 1916	ABI Illumina GS	U Wisconsin Sanger	AE014613.1	[6] [9]
Typhimurium LT2	ABI	WU	AE006468.1	[3]
Sequencing In Progress				
* Arizonae 05-0715 ATCC BAA-1577	454	WU	-	
* Diarizonae 61:1,v:1,5	ABI	WU	-	
* Diarizonae ATCC BAA-1579 05-0625 48:i:z	454	WU	-	
* Houtenae SARC13 ATCC BAA-1580 45:a:e,n,x	454	WU	-	
* Houtenae SARC14 ATCC BAA-1581 11:b:e,n,x	454	WU	-	

Sequences Completed	Method	Sequencing Center	GenBank Accession #	Citation
* Indica ATCC BAA-1576	ABI	WU	-	
* Indica ATCC BAA-1578	ABI	WU	-	
* Salamae ATCC BAA-1583 (05-0626) 47:b:1,5	454	WU	-	
* Salamae SARC3 58:d:z6	454	WU	-	
* Salmonella bongori 12149	ABI	Sanger	-	
4,[5],12:i:- CVM23701	ABI	TIGR	NZ_ABAO00000000	
Abortusovis SSM0041	454	WU	-	
Bovismorbificans 01-05481 PT13	454	WU	-	
Braenderup S-500	454	WU	-	
Brandenburg KMR12	ABI	Korea	-	
Dublin	ABI	U Illinois	-	
Enteritidis 48-0811	Illumina GS	WU	-	
Enteritidis LK5	ABI	U Illinois	-	
Enteritidis SARB17	454	WU	-	
Enteritidis SARB19	Illumina GS	WU	-	
Hadar	ABI	Sanger		
Hadar RI_05P066	ABI	TIGR/JCVI	NZ_ABFG00000000	
Heidelberg SL486	ABI	TIGR/JCVI	NZ_ABEL00000000	
Indiana KMR53	ABI	Korea	-	
Infantis	ABI	Sanger		
Infantis SARB27	454	WU	-	
Javiana GA_MM04042433	ABI	JCVI	NZ_ABEH00000000	
Kentucky CDC 191	ABI	JCVI	NZ_ABEI00000000	
Kentucky CVM29188	ABI	TIGR	NZ_ABAK00000000	
Miami ATCC BAA-1586 (02-3341)	454	WU	-	
Montevideo SARB30	454	WU	-	
Muenchen SARB32	454	WU	-	
Muenchen SARB34	454	WU	-	
Muenster ATCC BAA-1575 (0065-00)	ABI	WU	-	
Newport CVM36720	ABI	UMIGS	-	
Newport SL317	ABI	JCVI	NZ_ABEW00000000	
Panama KMR64	ABI	Korea	-	
Paratyphi A AKU_12601	ABI	Sanger	FM200053.1	
Paratyphi B ATCC BAA-1585	ABI	WU	-	
Paratyphi B SARB47	454	WU	-	
Paratyphi B tartrate (+) [Java] ATCC BAA-1584 (S-1241)	454	WU	-	
Paratyphi C RKS4594, SARB49	ABI Illumina GS	Peking University WU	- -	
Poona SGSC4934	454	WU	-	
Pullorum	ABI	U Illinois	-	

Sequences Completed	Method	Sequencing Center	GenBank Accession #	Citation
Saintpaul SARA23	ABI	TIGR	NZ_ABAM00000000	
Saintpaul SARA29	ABI	TIGR	NZ_ABAN00000000	
Schwarzengrund KMR78	ABI	Korea	-	
Schwarzengrund SL480	ABI	JCVI	NZ_ABEJ00000000	
Sendai 55-2461	454	WU	-	
Senftenberg SARB59	454	WU	-	
Stanley SARB60	454	WU	-	
Tennessee CDC07-0191	ABI	CDC	NZ_ACBF00000000	
Thompson SARB62	454	WU	-	
Typhi SGSC2661	Illumina GS	WU	-	
Typhi Ty21a	ABI	Naval Med. Res. Center	-	
Typhimurium 14028s	454	WU	-	
Typhimurium D23580	ABI	Sanger	-	
Typhimurium DT104	ABI	Sanger	-	
Typhimurium DT2	ABI	Sanger	-	
Typhimurium SL1344	ABI	Sanger	-	
Virchow SL491	ABI	TIGR/JCVI	NZ_ABFH00000000	
Weltevreden HI_N05-537	ABI	TIGR/JCVI	NZ_ABFF00000000	

* All sequences are from *Salmonella enterica* subspecies enterica except the marked strains, which are from other subspecies and *Salmonella bongori*.

WU = Washington University, St. Louis. **Sanger** = Wellcome Trust Sanger Institute. **JCVI** = J. Craig Venter Institute, **UMIGS** = University of Maryland

ABI – Sequencing of large fragments by the Sanger Method, using automated detection of an ABI DNA sequencer.

454- 454 Sequencing- (Roche) – sequencing of short fragments attached to beads by sequencing based on synthesis.

Illumina GS- Massively parallel sequencing of millions of fragments using proprietary terminator based sequencing chemistry.