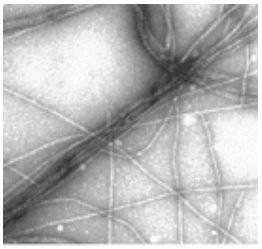
NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
NCBI Coffee Break

27 Oct 1999

Article reference: CB8.271099
Coffee Break archives



Proteins secreted by *Salmonella* interact with host cell actin and actin-binding proteins to assist *Salmonella* entry into host cells.

For more information and a complete figure, click here.

The *Salmonella* battle plan

$2.3 billion a year - that's the cost of *Salmonella* -related illness to the US economy, according to new data from the US Centers for Disease Control. The figure represents 1.4 million reported salmonellosis cases, including 600 deaths. Although the numbers are down from previous estimates, the human toll and dollar cost remain significant.

So how do people get sick from *Salmonella* infection? After eating contaminated food - often egg-based products - *Salmonella* gain access to the human intestine, where there is a low concentration of oxygen. This switches on a number of *Salmonella* genes that equip the bacteria with the proteins required to invade the cells on the surface of the intestine. A cocktail of these proteins is secreted and then translocated into the host epidermis cell. This stimulates the formation of a hollow in the host cell membrane, which grows deeper and progressively encloses the bacteria (a process known as pinocytosis). Eventually, the membrane pinches off to form a vacuole in which the bacterium resides. From this sheltered environment inside the intestine cell, the bacteria grow and divide, before spreading to other cells. When this happens, usually about two or three days after ingestion, the carrier experiences the symptoms of diarrhea and dehydration that are commonly associated with food poisoning.

All these rearrangements in the shape of the cell membrane mean that actin - the scaffolding of the cell - must undergo significant organizational change. Some of the secreted *Salmonella* proteins, whose job it is to manipulate the host's actin, have now been identified, for example, the 'Sips' (for *Salmonella* invasion proteins) - SipA, SipB, SipC, and SipD. Of these, SipA has recently been implicated in organizing actin filaments to facilitate *Salmonella* entry, by binding directly to actin itself or by binding to T-plastin, an actin-bundling protein. SipA lowers the concentration of actin monomers required for polymerization, allowing long filaments to be formed that extend the cell, and stabilizes those filaments so they can not break down again into actin monomers. But in reality, for the cell extensions to flex around the bacterium, actin filaments would need to breakdown and reform in an organized fashion. This suggests that there may be a regulatory mechanism that controls SipA action.

Shigella, which are found in the same family of bacteria as *Salmonella,* also invade intestine epithelial

Live PubMed searches
|| salmonella invasion || shigella entry || REVIEWS ||

BLAST tutorial
|| Find relatives of *Salmonella* SipA || VIEW ANALYSIS ||

Further NCBI resources
|| Taxonomy ||

cells on ingestion and causes diarrhea, fever, nausea and vomiting when they multiply and spread. Like *Salmonella, Shigella* secrete a collection of proteins that rearrange the actin cytoskeleton of the host. Among these, the proteins IpaA-D bear more than a passing resemblance to the Sips of *Salmonella,* suggesting a shared mechanism of entry between the two species. The further investigation of both bacterial invasion mechanisms may shed light on the evolutionary origins of this key virulence step, as well as provide targets for potential new therapies.

Comments?
Questions?
We would welcome feedback on NCBI's Coffee Break.
Email to: info@ncbi.nlm.nih.gov

Coffee Break
Article
BLAST tutorial

PubMed Links
Overview
Search PubMed

BLAST Links
Overview
BLAST page
Statistics course

The *Salmonella* battle plan

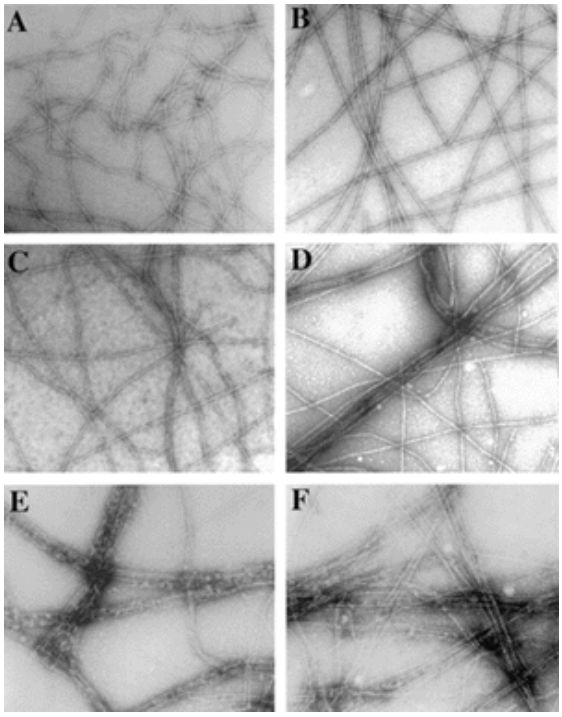


Figure 1. *Salmonella* SipA interacts with T-plastin to accelerate the bundling of actin filaments. The figure shows an electron micrograph of the effect of *Salmonella typhinurium* SipA on the bundling activity of T-plastin. **A and B** show actin with no T-plastin present; **C and D** show actin with low levels of T-plastin (0.125 micromolar), and **E and F** have a high level of T-plastin present (0.5 micromolar). **A, C and E** have no SipA present, whereas **B, D and F** have SipA present as 2.0 micromoles. The effect of SipA on actin bundling is especially clear in the pair **C and D**, where the concentration of T-plastin alone (**C**) is not sufficient to induce actin bundling. On addition of SipA (**D**), bundling occurs.

(Figure reproduced from Zhou, D., Mooseker, M.S. and Galan, J.E. (1999) An invasion-associated *Salmonella* protein modulates the actin-bundling activity of plastin *Proc. Natl Acad. Sci. USA* 96, 10176-10181 [PubMed])

Coffee Break
Article
Figure
Blast tutorial
Repeat tutorial

About Entrez
Entrez PubMed
Search
Overview
Help | FAQ

PubMed Services
Journal Browser
MeSH Browser
Single Citation Matcher
Batch Citation Matcher
Clinical Queries

Related Resources
Order Documents
Grateful Med
Consumer Health

The "salmonella AND food poisoning" search term gives rise to 1792 search results, listed in batches of 20 articles, starting with the most recent.

1 : Nylen G, et al. Related Articles
An international outbreak of Salmonella enteritidis associated with lasagne; lessons on the need for cross-national co-operation in investigating food-borne outbreaks.
Epidemiol Infect. 1999 Aug;123(1):31-5.
PMID: 10487639; UI: 99415580

2 : Okumara T, et al. Related Articles
[Sepsis due to Salmonella Oranienburg--a case report].
Kansenshogaku Zasshi. 1999 Aug;73(8):787-91. Japanese.
[MEDLINE record in process]
PMID: 10487026; UI: 99416494

3 : Yamada S, et al. Related Articles
[Recovery and its evaluation of Shigella bacilli or Salmonella from healthy food handlers in Tokyo (1961-1997)].
Kansenshogaku Zasshi. 1999 Aug;73(8):758-65. Japanese.
[MEDLINE record in process]
PMID: 10487021; UI: 99416489

4 : Bell RA, et al. Related Articles
The Abuela Project: safe cheese workshops to reduce the incidence of Salmonella typhimurium from consumption of raw-milk fresh cheese.
Am J Public Health. 1999 Sep;89(9):1421-4.
PMID: 10474564; UI: 99403812

5 : [No authors listed] Related Articles
From the Centers for Disease Control and Prevention. Outbreak of Salmonella serotype Muenchen infections associated with unpasteurized orange juice--United States and Canada, June 1999.
JAMA. 1999 Aug 25;282(8):726-8. No abstract available.
PMID: 10463698; UI: 99391339

6 : Hayes S, et al. Related Articles
Undercooked hens eggs remain a risk factor for sporadic Salmonella enteritidis infection.
Commun Dis Public Health. 1999 Jan;2(1):66-7.
PMID: 10462901; UI: 99392160

7 : Wilson D, et al. Related Articles
Common source outbreak of salmonellosis in a food factory.
Commun Dis Public Health. 1999 Jan;2(1):32-4.
PMID: 10462892; UI: 99392151

8 : Hedberg CW, et al. Related Articles
Outbreaks of salmonellosis associated with eating uncooked tomatoes: implications for public health. The Investigation Team.
Epidemiol Infect. 1999 Jun;122(3):385-93.
PMID: 10459640; UI: 99387396

9 : Evans MR, et al. Related Articles
An outbreak of Salmonella typhimurium DT170 associated with kebab meat and yogurt relish.
Epidemiol Infect. 1999 Jun;122(3):377-83.
PMID: 10459639; UI: 99387395

10 : Jahraus CD, et al. Related Articles
Clinical presentation and treatment of a Salmonella bredeney epidemic in Shelby County, Alabama.
South Med J. 1999 Aug;92(8):799-801.
PMID: 10456719; UI: 99383719

11 : [No authors listed] Related Articles
Outbreak of Salmonella serotype Muenchen infections associated with unpasteurized orange juice--United States and

Canada, June 1999.
MMWR Morb Mortal Wkly Rep. 1999 Jul 16;48(27):582-5.
PMID: 10428096; UI: 99355050

12 : Glosnicka R, et al. Related Articles
Comparison of two Salmonella enteritidis phage typing schemes.
Eur J Epidemiol. 1999 Apr;15(4):395-401.
PMID: 10414383; UI: 99341333

13 : Dieckmann H, et al. Related Articles
[Investigation of foodborne outbreak due to Salmonella infantis using epidemiological and microbiological methods].
Gesundheitswesen. 1999 May;61(5):241-7. German.
PMID: 10414018; UI: 99342423

14 : Neimann BJ, et al. Related Articles
[An outbreak of Salmonella enteritidis at the New Year celebration of the Copenhagen Medical Association].
Ugeskr Laeger. 1999 May 10;161(19):2803-6. Danish.
PMID: 10412319; UI: 99340816

15 : Molbak K. Related Articles
[Zoonotic bacterial gastroenteritis in Denmark].
Ugeskr Laeger. 1999 May 10;161(19):2796-9. Danish. No abstract available.
PMID: 10412316; UI: 99340813

16 : Lundstedt C, et al. Related Articles
[Is our food dangerous]?
Ugeskr Laeger. 1999 May 10;161(19):2789. Danish. No abstract available.
PMID: 10412313; UI: 99340810

17 : Lundstedt C, et al. Related Articles
[An editorial about an editorial].
Ugeskr Laeger. 1999 Jun 28;161(26):4022-3. Danish. No abstract available.
PMID: 10402943; UI: 99331354

18 : Przybylska A. Related Articles
[Foodborne infections and food poisoning in 1997].
Przegl Epidemiol. 1999;53(1-2):103-14. Polish.
PMID: 10402856; UI: 99331267

19 : Bellara SR, et al. Related Articles
Visualization and modelling of the thermal inactivation of bacteria in a model food.
Appl Environ Microbiol. 1999 Jul;65(7):3095-9.
PMID: 10388708; UI: 99318667

20 : Kusunoki J, et al. Related Articles
[Biochemical and molecular characterization of Salmonella ser. enteritidis phage type 1 isolated from food poisoning outbreaks in Tokyo].
Kansenshogaku Zasshi. 1999 May;73(5):437-44. Japanese.
PMID: 10386023; UI: 99313801

Revised: August 24, 1999.

NCBI

Tutorial Coffee Break

This is the query page for a BLAST search. The sequence of SipA is entered in the textbox below. The database to be searched can be selected from the following pull-down menu, as can the format that the sequence is submitted in. Click on the "Submit Query" button below the textbox to find similar sequences to this query in the database.

Database nr

The amino acid query sequence is filtered for low complexity regions by default.

Enter here your **amino acid sequence** as Sequence in FASTA format

```
VNVKPL SQL GVENVKTDTLQSDT
TVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNM
HSMHFDARETFLGVIRKALEPDTSTP
FPVRRAFDGLRAEILPNDTIKSAALKAQCSDIDKHPELKAKMETL
KEVITHHPQKEKLAEIALQFAREAG
LTRLKGETDYVLSNVLDGLIGDGSWRAGPAYESYLNKPGVDRVI
TTVDGLHMQR
```

Submit Query Reset

Please read about FASTA format description

The options below are for advanced users. In this example, the number of descriptions has been set to 50, with 10 alignments returned. If no parameters are changed or selected, then BLAST is run in default mode.

Advanced options for the BLAST server:

Expect 10 Filter default ☐ NCBI-gi ☑ Graphical Overview

Descriptions 50 Alignments 10

Expect value for inclusion in PSI-BLAST iteration 1 0.001

Other advanced options:

Submit Query Reset

Coffee Break
Article
Figure
Repeat tutorial
BLAST analysis

PubMed Links
Overview
Search PubMed

BLAST Links
Overview
BLAST page
Statistics course

Query= gi|1172128|gb|AAA86618.1| SipA (684 letters) **Database:** Non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR 414,351 sequences; 127,011,483 total letters

Below are the results from a PSI-BLAST (see Altschul et al., 1997) search of the non-redundant database using SipA as the query sequence.

The E value is a statistical measure of likelihood that the sequences listed below are truly similar to the query, rather than found by chance alone. The lower the E value, the greater the confidence that the protein found is a biologically significant match.

E-value threshold for inclusion in PSI-Blast iteration 1: 0.001
E-value threshold for inclusion in PSI-Blast iteration 2: 0.001

PSI-BLAST constructs a similarity matrix from the significant hits found in a BLAST search and uses this matrix to search the database in a second round of BLASTing. Each round of PSI-BLAST that uses the matrix to search the database is called an iteration.

Distribution of 22 Blast Hits on the Query Sequence

Mouse-over to show defline and

Color Key for Alignment Scores

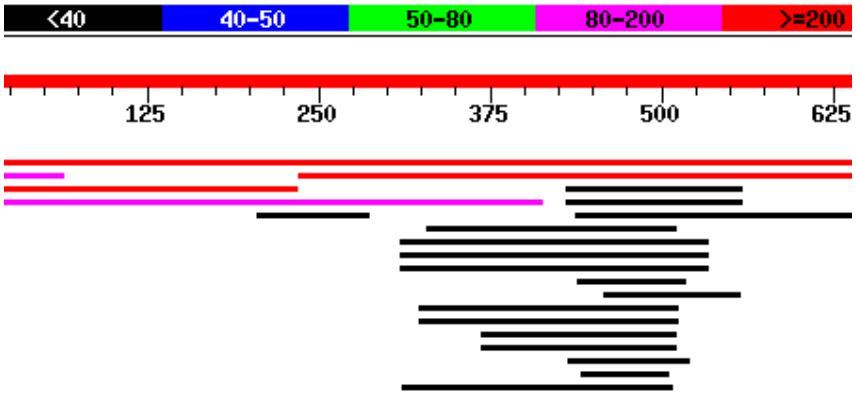


Figure legend

The above summary of the BLAST results represents proteins that match the query sequence as colored bars, with the most similar hit uppermost and appearing in red. Pink, green, blue and black bars follow, representing proteins in decreasing order of similarity. Hatched areas indicate a gap in similarity i.e., two or more distinct regions of similarity were found within the same protein hit. Moving the mouse over the bars will display the name of the matching protein found in the textbox above.

Click here to view alignments.
Click here to see the analysis of this BLAST result.

BLAST Alignments Coffee Break

About the alignments

Salmonella typhimurium SipA hits very few other proteins in the database that appear above the threshold e value of 0.001. The only proteins found to be significantly similar are SipA proteins from *Salmonella typhi* and IpaA from *Shigella flexneri.*

The match with the top score is to itself (reassuringly) - the lower the e-value or the higher the score, the higher the chance that the similarity is biologically relevant. Clicking on the scores will take you to the alignments for each of the top ten hits (as requested in our query). Clicking on the ID numbers on the left will display the sequence record at the NCBI site.

Back to BLAST result

Sequences with E-value BETTER than threshold

Sequences producing significant alignments:	Score (bits)	E Value
gi|1172128 (U40013) SipA [Salmonella typhimurium]	1227	0.0
emb|CAA63302| (X92546) sipA [Salmonella typhi]	767	0.0
emb|CAA57991| (X82670) sipA [Salmonella typhi] >gi|1584798|prf|...	394	e-108
gi|1117815 (U30491) SspA [Salmonella typhimurium]	108	2e-22
sp|P18010|IPAA_SHIFL 70 KD ANTIGEN >gi|2126189|pir||E31265 ipaA...	101	2e-20

Sequences with E-value WORSE than threshold

gi|1294769 (L35061) orfL12; this open reading frame codes for a...	39	0.082
gi|1353564 (U38906) ORF47 [Bacteriophage r1t]	37	0.42
sp|P48608|DIA_DROME DIAPHANOUS PROTEIN >gi|575927 (U11288) diap...	37	0.55
sp|P32107|YIBH_ECOLI HYPOTHETICAL 42.0 KD PROTEIN IN RHSA-MTLA ...	36	0.71
emb|CAA87091| (Z46970) secreted acid phosphatase 2 (SAP2) [Leis...	36	1.2
ref|NP_001933.1|PDSG1| desmoglein 1 >gi|416917|sp|Q02413|DSG1_H...	35	2.1
dbj|BAA23712| (AB007900) HH0452 cDNA clone for KIAA0440 has a 4...	35	2.1
gb|AAD12543.1| (AF090989) high-risk human papilloma viruses E6 ...	35	2.1
gb|AAD12544.1| (AF090990) high-risk human papilloma viruses E6 ...	35	2.1
dbj|BAA17666| (D90908) hypothetical protein [Synechocystis sp.]	34	3.6
sp|P44296|YH18_HAEIN HYPOTHETICAL PROTEIN HI1718 >gi|1074909|pi...	34	4.7
gi|2911243 (U67916) dentin sialophosphoprotein precursor [Mus m...	34	4.7
emb|CAA05208| (AJ002141) dentin sialophosphoprotein [Mus muscul...	34	4.7
emb|CAA71062.1| (Y09928) CTORF1365 [Staphylococcus aureus]	34	4.7
gb|AAD53452.1|AF137229_1 (AF137229) mixed lineage leukemia-like...	34	4.7
emb|CAB55329.1| (Y09927) Mrp protein [Staphylococcus aureus]	34	4.7
emb|CAA81224| (Z26314) STARP antigen [Plasmodium falciparum]	33	6.2

Alignments

```
gi|1172128 (U40013) SipA [Salmonella typhimurium]
          Length = 684

 Score = 1227 bits (3140), Expect = 0.0
 Identities = 621/684 (90%), Positives = 621/684 (90%)

Query: 1   MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGPQLEDFPALIKQ 60
           MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGPQLEDFPALIKQ
Sbjct: 1   MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGPQLEDFPALIKQ 60

Query: 61  ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG 120
           ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG
Sbjct: 61  ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG 120

Query: 121 AEYAAQIIKDGLKEKSAFGPWLPXXXXXXXXXXXXXXXXXDIIKNNTGGELSKLSTNLVM 180
           AEYAAQIIKDGLKEKSAFGPWLP                  DIIKNNTGGELSKLSTNLVM
Sbjct: 121 AEYAAQIIKDGLKEKSAFGPWLPETKKAEAKLENLEKQLLDIIKNNTGGELSKLSTNLVM 180

Query: 181 QEVMPYIASCIEHNFGCTLDPLTRSNLTHLVDKAAAKAVEALDMCPQKLTQEQGTSVGRE 240
           QEVMPYIASCIEHNFGCTLDPLTRSNLTHLVDKAAAKAVEALDMCPQKLTQEQGTSVGRE
Sbjct: 181 QEVMPYIASCIEHNFGCTLDPLTRSNLTHLVDKAAAKAVEALDMCPQKLTQEQGTSVGRE 240

Query: 241 ARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTGXXXXXXXXXX 300
           ARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTG
Sbjct: 241 ARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTGINININIDSS 300

Query: 301 XXXXXXXKHINNAEPVDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQRNGEXXXXXXXXXX 360
                 KHINNAEPVDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQRNGE
Sbjct: 301 NHSVDNSKHINNAEPVDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQRNGESHHSTNSSNV 360
```

```
Query: 361 XXXXXRVDSTTHQTETAHSASTGAIDHGIAGKIDVTAHATAEAVTNASSESKDGKXXXXX 420
                RVDSTTHQTETAHSASTGAIDHGIAGKIDVTAHATAEAVTNASSESKDGK
Sbjct: 361 SHSHSRVDSTTHQTETAHSASTGAIDHGIAGKIDVTAHATAEAVTNASSESKDGKVVTSE 420

Query: 421 XXXXXXXXXFDEVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVNVKPLASQLAGVEN 480
                   FDEVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVNVKPLASQLAGVEN
Sbjct: 421 KGTTGETTSFDEVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVNVKPLASQLAGVEN 480

Query: 481 VKTDTLQSDTTVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNMHSMHFD 540
           VKTDTLQSDTTVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNMHSMHFD
Sbjct: 481 VKTDTLQSDTTVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNMHSMHFD 540

Query: 541 ARETFLGVIRKALEPDTSTPFPVRRAFDGLRAEILPNDTIKSAALKAQCSDIDKHPELKA 600
           ARETFLGVIRKALEPDTSTPFPVRRAFDGLRAEILPNDTIKSAALKAQCSDIDKHPELKA
Sbjct: 541 ARETFLGVIRKALEPDTSTPFPVRRAFDGLRAEILPNDTIKSAALKAQCSDIDKHPELKA 600

Query: 601 KMETLKEVITHHPQKEKLAEIALQFAREAGLTRLKGETDYVLSNVLDGLIGDGSWRAGPA 660
           KMETLKEVITHHPQKEKLAEIALQFAREAGLTRLKGETDYVLSNVLDGLIGDGSWRAGPA
Sbjct: 601 KMETLKEVITHHPQKEKLAEIALQFAREAGLTRLKGETDYVLSNVLDGLIGDGSWRAGPA 660

Query: 661 YESYLNKPGVDRVITTVDGLHMQR 684
           YESYLNKPGVDRVITTVDGLHMQR
Sbjct: 661 YESYLNKPGVDRVITTVDGLHMQR 684
```

emb|CAA63302| (X92546) sipA [Salmonella typhi]
Length = 450

Score = 767 bits (1960), Expect = 0.0
Identities = 387/450 (86%), Positives = 394/450 (87%), Gaps = 1/450 (0%)

```
Query: 236 SVGREARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTGXXXXX 295
           SVGREARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTG
Sbjct: 1   SVGREARHLEMQTLIPLLLRNVFAQIPADKLPDPKIPEPAAGPVPDGGKKAEPTGININI 60

Query: 296 XXXXXXXXXXXXKHINNAEP-VDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQRNGEXXXX 354
                       KHINN+   VDN QRHIDNSNHDNSRKTIDNSRTFIDNSQR+GE
Sbjct: 61  NIDSSNHSVDNSKHINNSRSHVDNSQRHIDNSNHDNSRKTIDNSRTFIDNSQRHGESHHS 120

Query: 355 XXXXXXXXXXXRVDSTTHQTETAHSASTGAIDHGIAGKIDVTAHATAEAVTNASSESKDG 414
                     RVDSTTHQTETAHSASTG IDHGIAGKIDVTAHATAEAVTN+SSESKDG
Sbjct: 121 TNSSNVSHSHSRVDSTTHQTETAHSASTGTIDHGIAGKIDVTAHATAEAVTNSSSESKDG 180

Query: 415 KXXXXXXXXXXXXXXFDEVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVNVKPLASQ 474
           K               FDEVDGVTSKSIIGKP+QATVHGVDDNKQQSQTAEIVNVKPLASQ
Sbjct: 181 KVVTSEKGTTGETTSFDEVDGVTSKSIIGKPLQATVHGVDDNKQQSQTAEIVNVKPLASQ 240

Query: 475 LAGVENVKTDTLQSDTTVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNM 534
           LAGVENVK DTLQSD+TVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNM
Sbjct: 241 LAGVENVKIDTLQSDSTVITGNKAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNM 300

Query: 535 HSMHFDARETFLGVIRKALEPDTSTPFPVRRAFDGLRAEILPNDTIKSAALKAQCSDIDK 594
           HS+HF+ARE FLGVIRKALEPD STPFPVRRAFDGLR EILPNDTIKSAALKAQCSDIDK
Sbjct: 301 HSIHFNAREAFLGVIRKALEPDASTPFPVRRAFDGLRGEILPNDTIKSAALKAQCSDIDK 360

Query: 595 HPELKAKMETLKEVITHHPQKEKLAEIALQFAREAGLTRLKGETDYVLSNVLDGLIGDGS 654
           HPELKAKMETLKEVITHHPQKEKLAEIALQFAREAGLTR KGETDYVLSNVLDGLIGDGS
Sbjct: 361 HPELKAKMETLKEVITHHPQKEKLAEIALQFAREAGLTRQKGETDYVLSNVLDGLIGDGS 420

Query: 655 WRAGPAYESYLNKPGVDRVITTVDGLHMQR 684
           WRAGPAYESYLNKPGVDRVITTVDGLHMQR
Sbjct: 421 WRAGPAYESYLNKPGVDRVITTVDGLHMQR 450
```

emb|CAA57991| (X82670) sipA [Salmonella typhi] >gi|1584798|prf||2123407E sipA
gene [Salmonella typhi]
Length = 235

Score = 394 bits (1001), Expect = e-108
Identities = 204/235 (86%), Positives = 206/235 (86%)

```
Query: 1   MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGPQLEDFPALIKQ 60
           MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVR    +    EIKG QLEDFPALIKQ
Sbjct: 1   MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRESATATLSGEIKGQQLEDFPALIKQ 60

Query: 61  ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG 120
           ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG
Sbjct: 61  ASLDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVG 120

Query: 121 AEYAAQIIKDGLKEKSAFGPWLPXXXXXXXXXXXXXXXXXDIIKNNTGGELSKLSTNLVM 180
           AEY AQIIKDGLKEKSAFGPWLP                 DIIKNNTGGELSKLSTNLVM
Sbjct: 121 AEYTAQIIKDGLKEKSAFGPWLPETKKAEAKLENLEKQLLDIIKNNTGGELSKLSTNLVM 180

Query: 181 QEVMPYIASCIEHNFGCTLDPLTRSNLTHLVDKAAAKAVEALDMCPQKLTQEQGT 235
           QEVMPYIASCIEHNFGCTLDPLTRS+LT LVDKAAAKAVEALDMC QKLTQEQGT
Sbjct: 181 QEVMPYIASCIEHNFGCTLDPLTRSSLTQLVDKAAAKAVEALDMCHQKLTQEQGT 235
```

```
 gi|1117815 (U30491) SspA [Salmonella typhimurium]
          Length = 65

 Score =  108 bits (266), Expect = 2e-22
 Identities = 56/65 (86%), Positives = 57/65 (87%)

Query: 1  MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGPQLEDFPALIKQ 60
          MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVR    +    EIKGPQLEDFPALIKQ
Sbjct: 1  MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRESATATLSGEIKGPQLEDFPALIKQ 60

Query: 61 ASLDA 65
          ASLDA
Sbjct: 61 ASLDA 65


 sp|P18010|IPAA_SHIFL 70 KD ANTIGEN >gi|2126189|pir||E31265 ipaA protein - Shigella
           flexneri plasmid >gi|47031|emb|CAA35624| (X17628)
           invasion plasmid antigen (AA 1-633( [Shigella flexneri]
           Length = 633

 Score =  101 bits (249), Expect = 2e-20
 Identities = 99/423 (23%), Positives = 163/423 (38%), Gaps = 49/423 (11%)

Query: 7   TQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVREIKGP----QLEDFPALIKQAS 62
           TQ P  +    +   T+ + L + +S + +   S     +K P    + E+F     Q
Sbjct: 7   TQAPTFLYKATSPSSTEYSELKSKISDIHSSQTS-----LKTPASVSEKENFATSFNQKC 61

Query: 63  LDALFKCGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVGAE 122
           LD LF      + L+ +++NS N   K  I+EF+ +  S ++    + E +  L K+ A+
Sbjct: 62  LDFLFSSSGKEDVLRSIYSNSMNAYAKSEILEFSNVLYSLVHQNGLNFENEKGLQKIVAQ 121

Query: 123 YAAQIIKDGLKEKSAFGPWLPXXXXXXXXXXXXXXXXXDIIKNNTGGELSKLSTNLVMQE 182
           Y+  IIKD L + SAFGPW                   + + +T GE   L   L+  E
Sbjct: 122 YSELIIKDKLSQDSAFGPWSAKNKKLHQLRQNIEHRLALLAQQHTSGEALSLGQKLLNTE 181

Query: 183 VMPYIASCIEHNFGCTLDPLTRSNLTHLVDKAAAKAVEALDMCPQKLTQEQGTSVGREAR 242
           V  +I + I      + + ++   L  LVD  A  A ++L    +     +G  +G+ +R
Sbjct: 182 VSSFIKNNILAELKLSNETVSSLKLDDLVDAQAKLAFDSLRNQRKNTIDSKGFGIGKLSR 241

Query: 243 HLEMQTLIPLLLRNVFAQI---------PADKLPDPKIPEPAAGPVPDGGKKAEPTGXXX 293
            L    + P LLR V   I           D LP P    P  GP P   +K
Sbjct: 242 DLNTVAVFPELLRKVLNDILEDIKDSHPIQDGLPTPPEDMPDGGPTPGANEK-------- 293

Query: 294 XXXXXXXXXXXXXXKHINNAEPVDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQRNGEXXX 353
                              ++PV +   HI+N N     +  DN R + ++   N E
Sbjct: 294 ------------------TSQPVIH--YHINNDNRTYDNRVFDN-RVYDNSYHENPENDA 332

Query: 354 XXXXXXXXXXXXR-VDSTTHQTETAHSASTGAIDHGIAGKIDVTAHATA-EAVTNASSES 411
                       R  +S  +         T  + H I+  +   A+ +A E V N + ++
Sbjct: 333 QSPTSQTNDLLSRNGNSLLNPQRALVQKVTSVLPHSISDTVQTFANNSALEKVFNHTPDN 392

Query: 412 KDG 414
            DG
Sbjct: 393 SDG 395


 gi|1294769 (L35061) orfL12; this open reading frame codes for a putative
           protein sharing 27-91 % homology with the putative
           protein encoded by orf47 of the lactococcal
           bacteriophage r1t; putative [Bacteriophage phi-41]
           Length = 586

 Score = 39.5 bits (90), Expect = 0.082
 Identities = 36/143 (25%), Positives = 63/143 (43%), Gaps = 17/143 (11%)

Query: 432 EVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVN-VKPLASQLAGVEN---------V 481
           EV G+T ++++   +++ +  +  N  QSQ +E++N V  L S ++ + N
Sbjct: 98  EVFGITPENLV--TIKSLLDAIASNASQSQVSELINSVSTLTSNISLMSNGDYSPKANQT 155

Query: 482 KTDTLQ---SDTTVITGNKAGTTDNDNSQ--TDKTGPFSGLKFKQNSFLSTVPSVTNMHS 536
             ++LQ   +D +     KA  TD DN Q   DK G     K +Q+    T  +VT +
Sbjct: 156 DLESLQHTVNDHSATISTKANQTDLDNLQATVDKQGVAISTKAEQSELSITNKNVTTVQE 215

Query: 537 MHFDARETFLGVIRKALEPDTST 559
               A       + KA E   ++
Sbjct: 216 TAKQAESEAKNAMAKATEAQANS 238


 gi|1353564 (U38906) ORF47 [Bacteriophage r1t]
           Length = 667

 Score = 37.1 bits (84), Expect = 0.42
 Identities = 34/143 (23%), Positives = 62/143 (42%), Gaps = 17/143 (11%)

Query: 432 EVDGVTSKSIIGKPVQATVHGVDDNKQQSQTAEIVN-VKPLASQLAGVEN---------V 481
           E  G++ ++++   +++ +  +  N  +S+  E++N VK L S ++ + N
Sbjct: 192 EAFGISPENLV--TIKSLLEAIASNASESEVVELINSVKVLTSNISLMSNGDYSPKANQT 249

Query: 482 KTDTLQ---SDTTVITGNKAGTTDNDNSQ--TDKTGPFSGLKFKQNSFLSTVPSVTNMHS 536
```

```
            ++LQ   +D +     KA  TD DN Q   DK G     K +Q+  L T  +VT
Sbjct: 250 DLESLQHTVNDQSATISTKANQTDLDNLQATVDKQGIAISTKAEQSELLITNKNVTTAQE 309

Query: 537 MHFDARETFLGVIRKALEPDTST 559
               A       + KA E   ++
Sbjct: 310 TAKQAESEAKNAMAKATEAQANS 332


 sp|P48608|DIA_DROME DIAPHANOUS PROTEIN >gi|575927 (U11288) diaphanous protein
           [Drosophila melanogaster]
           Length = 1091

 Score = 36.7 bits (83), Expect = 0.55
 Identities = 25/82 (30%), Positives = 38/82 (45%), Gaps = 3/82 (3%)

Query: 207 LTHLVDKAAAKAVEALDMCPQKLTQEQGTSVGREARHLEMQTLIPLLLRNVFAQIPADKL 266
           L  +V+KA AK  +  +   +K+ Q +      EA+   ++  + L+  N  A    +KL
Sbjct: 447 LDDIVEKAKAKESKRSEEYEKKIEQLESAKQEAEAKAAHLEEKVKLMEANGVAAPSPNKL 506

Query: 267 PDPKIPEPAAGPVPDGGKKAEP 288
           P   IP P   P P GG  A P
Sbjct: 507 PKVNIPMP---PPPPGGGGAPP 525


 sp|P32107|YIBH_ECOLI HYPOTHETICAL 42.0 KD PROTEIN IN RHSA-MTLA INTERGENIC REGION
           >gi|1073455|pir||S47818 hypothetical protein f378b -
           Escherichia coli >gi|466735 (U00039) alternate gene name
           yibH [Escherichia coli] >gi|1790024 (AE000437) putative
           membrane protein [Escherichia coli] >gi|1857028 (L19044)
           unknown [Escherichia coli]
           Length = 378

 Score = 36.4 bits (82), Expect = 0.71
 Identities = 61/280 (21%), Positives = 110/280 (38%), Gaps = 49/280 (17%)

Query: 439 KSIIGKPVQATVHGV-----DDNKQQSQTAEIV-NVKPLASQLAGVENVKTDTLQSDTTV 492
           K++I  P+   V G+     D N Q  Q  E++  + P+  Q       + D LQ+D
Sbjct: 59  KAVIAIPITPQVTGIVTEVTDKNNQLIQKGEVLFKLDPVRYQ------ARVDRLQADLMT 112

Query: 493 ITGN----KAGTTDNDNSQTDKTGPFSGLKFKQNSFLSTVPSVTNMHSMHF--DARETFL 546
            T N    +A  T+   + T  +     L      +L    +  N  S     DAR+ FL
Sbjct: 113 ATHNIKTLRAQLTEAQANTTQVSAERDRLFKNYQRYLKGSQAAVNPFSERDIDDARQNFL 172

Query: 547 G---VIRKALEPDTSTPFPVRRAFDGLRAEILP------------NDTIKSAALKAQCSD 591
               +++ ++         +    +G +++I+               T+  A      +
Sbjct: 173 AQDALVKGSVAEQAQIQSQLDSMVNGEQSQIVSLRAQLTEAKYNLEQTVIRAPSNGYVTQ 232

Query: 592 IDKHPELKAKMETLKEVITHHPQKEKLAEIALQFAREAGLTRLK-GETDYVLSNVLDG-- 648
           +   P   A    L+ V+   P++++  +I  QF R+  L RLK G+   V+ N L G
Sbjct: 233 VLIRPGTYAAALPLRPVMVFIPEQKR--QIVAQF-RQNSLLRLKPGDDAEVVFNALPGQV 289

Query: 649 ----------LIGDGSWRAGPAYESYLNKPGVDRVITTVD 678
                     ++  GS++A    +S    PG D V+ T++
Sbjct: 290 FHGKLTSILPVVPGGSYQAQGVLQSLTVVPGTDGVLGTIE 329


 emb|CAA87091| (Z46970) secreted acid phosphatase 2 (SAP2) [Leishmania mexicana]
           Length = 888

 Score = 35.6 bits (80), Expect = 1.2
 Identities = 32/182 (17%), Positives = 57/182 (30%), Gaps = 9/182 (4%)

Query: 330 NSRKTIDNSRTFIDNSQRNGEXXXXXXXXXXXXXXXRVDSTTHQTETAHSASTGAIDHGI 389
           +S  T   S     +S   G                  S+   T ++   +T +
Sbjct: 658 SSEGTATTSSDATTSSSSEGTTSSSSDATTSSSEGTATTSSDATTSSSSEGTTSSRSDAT 717

Query: 390 AGKIDVTAHATAEAVTNASSESKDGKXXXXXXXXXXXXXXXFDEVDGVTSKSIIGKPVQAT 449
               + TA  +++A T++SSE                     D  TS S  G
Sbjct: 718 TSSSEGTATTSSDATTSSSSEGTTSSSSDATTSSSEGTATTSS-DATTSSSSEGT----- 771

Query: 450 VHGVDDNKQQSQTAEIVNVKPLASQLAGVENVKTDTLQSDTTVITGNKAGTTDNDNSQTD 509
                 +   + T+  V     +S+     +    T  SD T  + +  GTT + +S +
Sbjct: 772 ---TSSSSDATTTSSDVTTTSSSSEGTATSSSDATTTSSDVTTTSSSSEGTTTSSSSSSS 828

Query: 510 KT 511
           K+
Sbjct: 829 KS 830


CPU time:    62.31 user secs.       0.71 sys. secs         63.02 total secs.

  Database: Non-redundant GenBank CDS
  translations+PDB+SwissProt+SPupdate+PIR
    Posted date:  Sep 30, 1999  5:53 PM
  Number of letters in database: 127,011,483
  Number of sequences in database:  414,351

Lambda     K      H
   0.312    0.129    0.361
```

```
Gapped
Lambda     K      H
   0.270   0.0470    0.230


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 148004274
Number of Sequences: 414351
Number of extensions: 5573570
Number of successful extensions: 14523
Number of sequences better than 10.0: 22
Number of HSP's better than 10.0 without gapping: 6
Number of HSP's successfully gapped in prelim test: 16
Number of HSP's that attempted gapping in prelim test: 14365
Number of HSP's gapped (non-prelim): 41
length of query: 684
length of database: 127011483
effective HSP length: 63
effective length of query: 621
effective length of database: 100907370
effective search space: 62663476770
effective search space used: 62663476770
T: 11
A: 40
X1: 16 ( 7.2 bits)
X2: 38 (14.8 bits)
X3: 64 (24.9 bits)
S1: 42 (21.9 bits)
S2: 73 (32.8 bits)
```

Analysis Coffee Break

The *Salmonella* battle plan

The SipA proteins from *S. typhinurium* and *S. typhi* are almost identical, and significantly similar to *Shigella* IpaA protein.

```
S. typhi SipA        MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRESATATLSGGIKG----QQLEDFPALIKQASLDALFK
S. typhinurium SipA  MVTSVRTQPPVIMPGMQTEIKTQATNLAANLSAVRAKCHSDAVRGIKGP----QLEDFPALIKQASLDALFK
Shigella IpaA        MHMVNNTQAPTFLYKATSPSSTEYSELKSKISDIHSSQTS-----LKTPASVSEKENFATSFNQKCLDFLFS

S. typhi SipA        CGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVGAEYTAQIIKDGLKEKSAFGP
S. typhinurium SipA  CGKDAEALKEVFTNSNNVAGKKAIMEFAGLFRSALNATSDSPEAKTLLMKVGAEYAAQIIKDGLKEKSAFGP
Shigella IpaA        SSGKEDVLRSIYSNSMNAYAKSEILEFSNVLYSLVHQNGLNFENEKGLQKIVAQYSELIIKDKLSQDSAFGP

S. typhi SipA        WLPETKKAEAKLENLEKQLLDIIKNNTGGELSKLSTNLVMQEVMPYIASCIEHNFGCTLDPLTRSSLTQLVD
S. typhinurium SipA  WLPETKKAEAKLENLEKQLLDIIKNNTGGELSKLSTNLVMQEVMPYIASCIEHNFGCTLDPLTRSMLTHLVD
Shigella IpaA        WSAKNKKLHQLRQNIEHRLALLAQQHTSGEALSLGQKLLNTEVSSFIKNNILAELKLSNETVSSLKLDDLVD

                                                •
S. typhi SipA        KAAAKAVEALDMQHQKLTQEQGTSVGREARHLEMQTLIPLLLRNVFAQI--------P-ADKLPDPKIPEPA
S. typhinurium SipA  KAAAKAVEALDMQPQKLTQEQGTSVGREARHLEMQTLIPLLLRNVFAQI--------P-ADKLPDPKIPEPA
Shigella IpaA        AQAKLAFDSLRNQRKNTIDSKGFGIGKLSRDLNTVAVFPELLRKVLNDILEDIKDSHPIQDGLPTPPEDMPD

S. typhi SipA        AGPVPDGGKKAEPTGINININIDSSNHSVDNSKHINNSRSHVDNSQRHIDNSNHDNSRKTIDNSRTFIDNSQ
S. typhinurium SipA  AGPVPDGGKKAEPTGINININIDSSNHSVDNSKHINN-AEPVDNGQRHIDNSNHDNSRKTIDNSRTFIDNSQ
Shigella IpaA        GGPTPGANEK--------------TSQPV-IHYHINNDNRTYDNR--VFDNRVYDNSYHENPENDAQSPTSQ

S. typhi SipA        RHGESHHSTNSSNVSHSHSRVDS--TTHQTETAHSASTGNIDHGIAGKIDVTAHATAEAVTNSSSESKDG +270
S. typhinurium SipA  RMGESHHSTNSSNVSHSHSRVDS--TTHQTETAHSASTGAIDHGIAGKIDVTAHATAEAVTNASSESKDG +270
Shigella IpaA        --------TN-----DLLSRNGNSLLNPQ-RALVQKVTSVLPHSISDTVQTFANNSALEKVFNHTPDNDG +238
```

The above alignment shows the sequences of *Salmonella typhinurium* and *Salmonella typhi* SipA with *Shigella flexneri* IpaA protein. The SipA proteins are almost identical, with the exception of a few amino acid differences, indicated by the turquoise boxes. Amino acids common to all three proteins are in red. The *S. typhi* sequence shown here is made up of two separate GenBank records that each aligned to different parts of the complete *S. typhinurium* sequence (blue background). This is indicated by the yellow (accession number CAA57991) and green (accession number CAA63302) highlighting behind the two sequences, and their junction is shown by the black dot. The numbers at the end of the sequences indicate the number of amino acids not included in the alignment: while the *S. typhinurium* and *S. typhi* SipA proteins remain highly similar in this region, there is no similarity between the SipAs and IpaA.

Back to BLAST result

Salmonella typhinurium SipA was used to search the non-redundant protein sequence database using the PSI-BLAST program [2] with standard parameters. Sequences included in the multiple alignment were selected from output from the first iteration. The multiple sequence alignment was constructed using ClustalW [3].

[1] Zhou, D., Mooseker, M.S. and Galan, J.E. (1999) An invasion-associated Salmonella protein modulates the actin-bundling activity of plastin. *Proc. Natl Acad. Sci. USA* 96, 10176-10181
[2] Altschul, S. F. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-402
[3] Higgins, D. G., Thompson, J. D. and Gibson, T. J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 266, 383-402