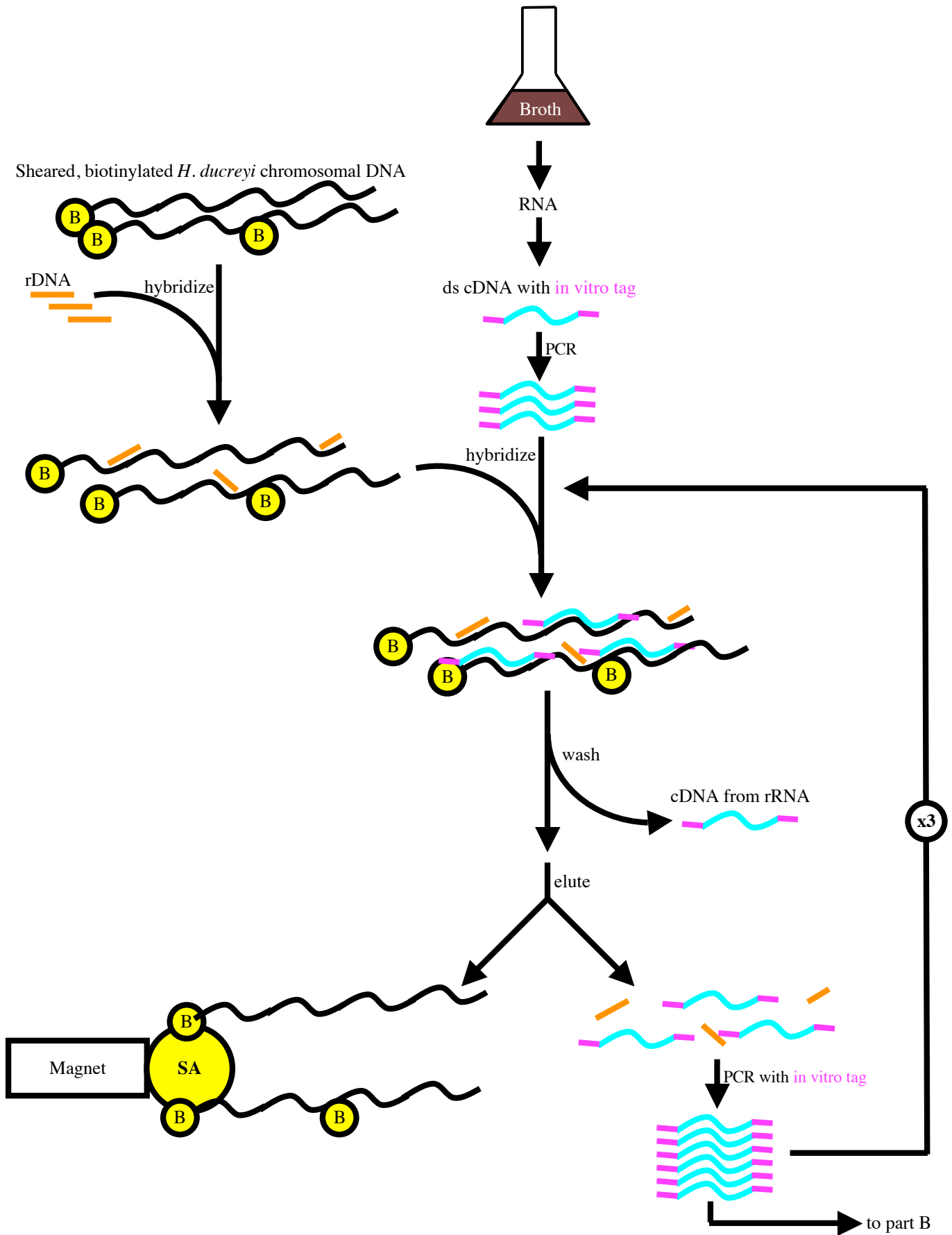
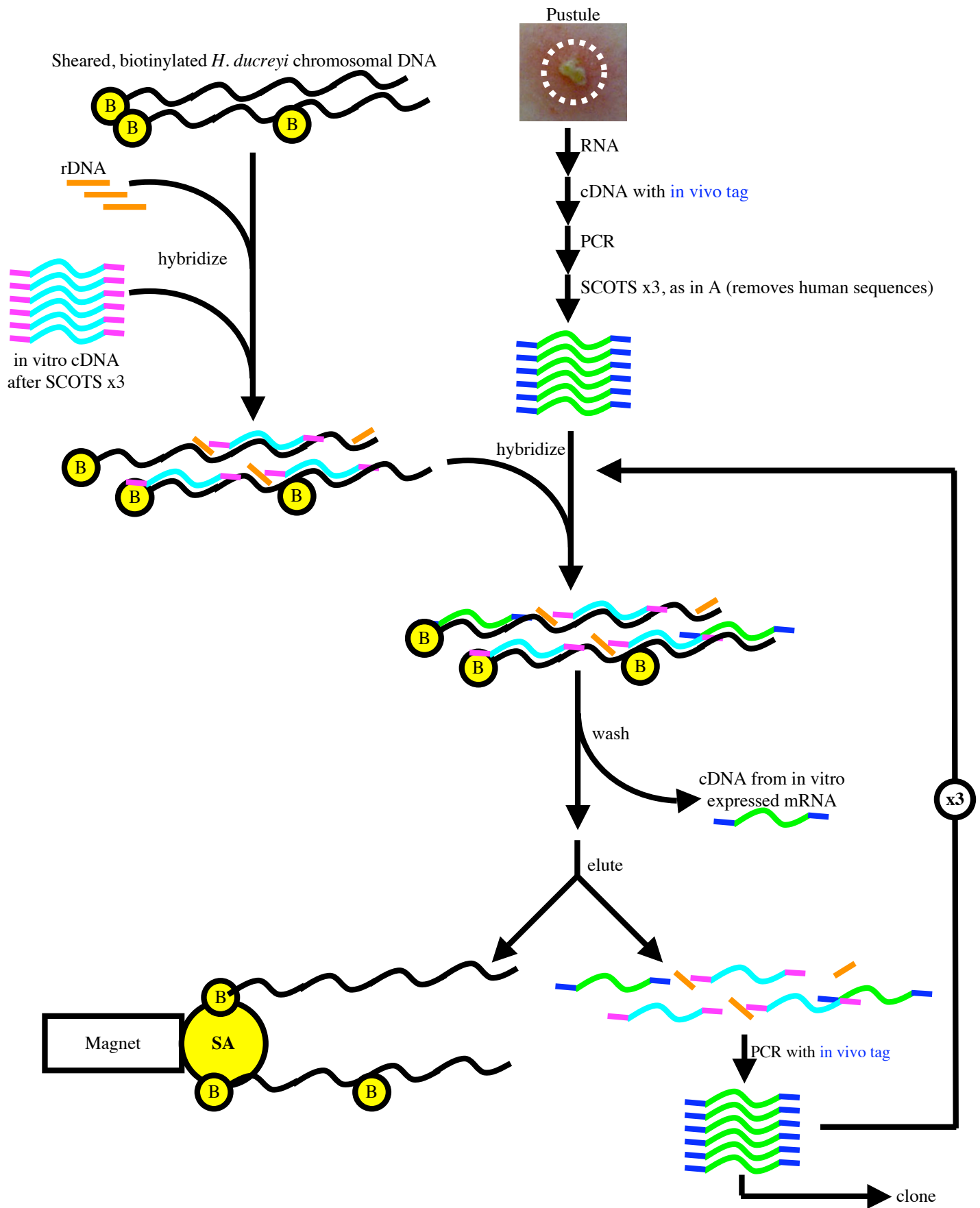


# A. SCOTS performed on in vitro-derived cDNA



## B. Competitive SCOTS performed on in vivo-derived cDNA



## Figure 1S Legend.

SCOTS Procedure. **(A)** Non-competitive SCOTS. Total RNA is isolated from broth-grown bacteria (shown) and converted to double-stranded cDNA (light blue) with an in vitro-specific tag (pink) for PCR amplification. The amplified cDNA is hybridized with sheared, biotinylated *H. ducreyi* chromosomal DNA (black) that has been prehybridized with rDNA (orange). cDNA representing genes transcribed in vitro bind to the biotinylated chromosomal DNA, while rRNA-derived sequences are blocked from hybridization and washed off. Then, mRNA-derived cDNA is eluted and PCR-amplified, and the hybridization repeated a total of three times. **(B)** Competitive SCOTS. A pustule from the human model of *H. ducreyi* infection is excisionally biopsied (dashed white ring). Total RNA is isolated, converted to double-stranded cDNA (green) with in vivo-specific tag (dark blue) for PCR amplification. The cDNA is subjected to 3 rounds of non-competitive SCOTS as described in (A) to remove human and rRNA derived sequences. After the final PCR amplification, in vivo-derived cDNA is hybridized to sheared, biotinylated *H. ducreyi* chromosomal DNA that has been pre-hybridized with rDNA and in vitro derived cDNA from (A). The in vitro derived cDNA hybridizes to genes transcribed in vitro, and only genes transcribed in vivo are available for subsequent hybridization. After washing to remove nonhybridized cDNA, all hybridized cDNA is eluted and PCR amplified with the in vivo specific tag to preferentially amplify in vivo transcribed sequences. The PCR amplified cDNA is then subjected to two additional rounds of hybridization, after which the final cDNA pool is cloned for sequencing. See section 4 for details.

1 Table S1. Tissue Samples Used in This Study

2

Subject #	Gender <sup>a</sup>	# Days Infected	Tissue Used for:	Reference
238	F	6	SCOTS	(Janowicz <i>et al.</i> , 2004)
240	F	6	SCOTS	(Janowicz <i>et al.</i> , 2004)
250	F	7	SCOTS	(Fulcher <i>et al.</i> , 2006)
142	M	7	RT-PCR	(Throm & Spinola, 2001)
160	M	6	RT-PCR	(Bong <i>et al.</i> , 2001)
164	M	7	RT-PCR	(Throm & Spinola, 2001)
231	M	6	RT-PCR	(Humphreys <i>et al.</i> , 2005)
232	M	6	RT-PCR	(Humphreys <i>et al.</i> , 2005)
249	M	8	RT-PCR	(Fulcher <i>et al.</i> , 2006)
252	F	9	RT-PCR	(Fulcher <i>et al.</i> , 2006)
276	M	6	RT-PCR	(Janowicz <i>et al.</i> , 2006a)
281	F	6	RT-PCR	(Janowicz <i>et al.</i> , 2006a)

3

4 <sup>a</sup> F, female; M, male.

5

8

## Supplementary Data Table S2

Functional Category	Gene ID	Gene Name or Homolog	Putative Function
Adherence-related	HD1304	<i>tadA</i> , tight adherence protein A	possible adhesin or involved in type IV secretion
Adherence-related	HD1895	putative adhesin processing HmwC-like protein	
Bacteriophage Genes	HD0527	possible tail length tape measure protein	morphogenesis, determines tail length
Bacteriophage Genes	HD0518	<i>gpI</i> , bacteriophage Mu I protein GP32	protease involved in virion morphogenesis
Bacteriophage Genes	HD0121	Mu-like phage gp27	
Bacteriophage Genes	HD0124	Mu-like phage gp29	
Bacteriophage Genes	HD0132	possible Mu-like phage protein	
Bacteriophage Genes	HD0143	possible bacteriophage TM4 gp14 protein homolog	
Bacteriophage Genes	HD0514	possible Mu-like phage protein	
Biosynthesis/Metabolism	HD1852	<i>nanE</i> , possible N-acetylmannosamine-6-P epimerase	A-acetylmannosamine, sialic acid utilization pathways
Biosynthesis/Metabolism	HD0566	<i>argE</i> , acetylornithine deacetylase	amino acid synthesis: arginine biosynthesis
Biosynthesis/Metabolism	HD0890	<i>argC</i> , N-acetyl-gamma-glutamyl-phosphate reductase	amino acid synthesis: arginine biosynthesis
<b>Biosynthesis/Metabolism</b>	<b>HD1384</b>	<b><i>argA</i>, amino-acid acetyltransferase</b>	<b>amino acid synthesis: arginine biosynthesis</b>
<b>Biosynthesis/Metabolism</b>	<b>HD0233</b>	<b><i>carB</i>, carbamoyl-phosphate synthase, large subunit</b>	<b>amino acid synthesis: arginine, pyrimidine biosynthesis</b>
Biosynthesis/Metabolism	HD1383	<i>aroA</i> , 3-phosphoshikimate 1-carboxyvinyltransferase	amino acid synthesis: aromatic a.a. biosynthesis
Biosynthesis/Metabolism	HD1709	<i>glnE</i> , glutamate ammonia ligase adenyllyltransferase	amino acid synthesis: modification of glutamine synthetase, allosteric control of glutamine
Biosynthesis/Metabolism	HD0372	<i>ansB</i> , L-asparaginase II	breaks down asparagine to aspartate, anaerobic induction in <i>E. coli</i>
<b>Biosynthesis/Metabolism</b>	<b>HD2020</b>	<b><i>mrDA</i>, penicillin binding protein 2</b>	<b>cell wall structure: peptidoglycan degradation</b>
Biosynthesis/Metabolism	HD0241	<i>ftsI</i> , penicillin binding protein 3	cell wall structure: peptidoglycan synthesis
Biosynthesis/Metabolism	HD0426	<i>mrcA</i> , penicillin binding protein 1A	cell wall structure: peptidoglycan synthesis
<b>Biosynthesis/Metabolism</b>	<b>HD0823</b>	<b><i>murC</i>, UDP-N-acetylmuramate--alanine ligase</b>	<b>cell wall structure: peptidoglycan synthesis</b>
Biosynthesis/Metabolism	HD1275	<i>mepA</i> , penicillin-insensitive murein endopeptidase A	cell wall structure: peptidoglycan synthesis
Biosynthesis/Metabolism	HD0729	<i>dat</i> , diaminobutyrate--pyruvate aminotransferase	central intermediary metabolism: amine, polyamine, ectoine biosynthesis
Biosynthesis/Metabolism	HD1845	<i>nagA</i> , N-acetylglucosamine-6-phosphate deacetylase	central intermediary metabolism: N-acetylglucosamine utilization pathway
Biosynthesis/Metabolism	HD1624	<i>aceF</i> , dihydrolipoamide acetyltransferase	central intermediary metabolism: part of pyruvate dehydrogenase complex
<b>Biosynthesis/Metabolism</b>	<b>HD1082</b>	<b><i>iscS</i>, cysteine desulfurase</b>	<b>cofactor biosynthesis: iron-sulfur cluster formation, removes sulfur from Cys</b>
Biosynthesis/Metabolism	HD0214	<i>pdxH</i> , pyridoxamine-5'-phosphate oxidase	cofactor biosynthesis: pyridoxine (vitamin B6) production
<b>Biosynthesis/Metabolism</b>	<b>HD1593</b>	<b><i>pdxS</i>, putative pyridoxine biosynthesis protein</b>	<b>cofactor biosynthesis: pyridoxine (vitamin B6) production</b>
Biosynthesis/Metabolism	HD0333	<i>trxB</i> , thioredoxin reductase	cofactor biosynthesis: reduces thioredoxin, protection from oxidative stress
Biosynthesis/Metabolism	HD0273	<i>ribF</i> , riboflavin biosynthesis protein	cofactor biosynthesis: riboflavin biosynthesis
Biosynthesis/Metabolism	HD1806	<i>thiI</i> , thiamine biosynthesis protein	cofactor biosynthesis: thiamine biosynthesis
Biosynthesis/Metabolism	HD0705	<i>ispB</i> , octaprenyl-diphosphate synthase	cofactor biosynthesis: ubiquinone, menaquinone synthesis
<b>Biosynthesis/Metabolism</b>	<b>HD1457</b>	<b><i>pta</i>, phosphate acetyltransferase</b>	<b>conversion of acetate to acetyl CoA</b>
Biosynthesis/Metabolism	HD1842	<i>wecC</i> , UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase	ECA biosynthesis
Biosynthesis/Metabolism	HD1844	<i>wecA</i> , undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase	ECA biosynthesis, O-Ag biosynthesis
Biosynthesis/Metabolism	HD0707	<i>fabD</i> , malonyl CoA-acyl carrier protein transacylase	fatty acid and phospholipid metabolism: fatty acid biosynthesis
<b>Biosynthesis/Metabolism</b>	<b>HD0634</b>	<b><i>accB</i>, biotin carboxyl carrier protein of acetyl-CoA carboxylase</b>	<b>fatty acid and phospholipid metabolism: long chain fatty acid biosynthesis</b>
Biosynthesis/Metabolism	HD0829	<i>galE</i> , UDP-glucose-4-epimerase	galactose metabolism: controls internal galactose concentration
Biosynthesis/Metabolism	HD0465	<i>pfkA</i> , phosphofructokinase	glycolysis/gluconeogenesis: controls glycolysis
Biosynthesis/Metabolism	HD0990	<i>pfkB</i> , formate acetyltransferase	glycolysis/gluconeogenesis: nonoxidative glucose metabolism
<b>Biosynthesis/Metabolism</b>	<b>HD1331</b>	<b><i>pckA</i>, phosphoenolpyruvate carboxykinase</b>	<b>glycolysis/gluconeogenesis: rate-limiting gluconeogenesis</b>
Biosynthesis/Metabolism	HD0217	<i>lpxK</i> , tetraacyldisaccharide 4' kinase	LOS biosynthesis: Lipid A biosynthesis
Biosynthesis/Metabolism	HD1101	<i>kdkA</i> , 3-deoxy-D-mannooctulosonic acid kinase	LOS biosynthesis: LOS core biosynthesis
Biosynthesis/Metabolism	HD1182	<i>waaE</i> , ADP-heptose synthase	LOS biosynthesis: LOS core biosynthesis
Biosynthesis/Metabolism	HD1929	<i>rpe</i> , D-ribulose-phosphate-3 epimerase	pentose-phosphate pathway: xylulose synthesis
Biosynthesis/Metabolism	HD0548	<i>purT</i> , phosphoribosylglycinamide formyltransferase 2	purine biosynthesis
Biosynthesis/Metabolism	HD1247	<i>maeA</i> , malate oxidoreductase	TCA cycle, catalyzes malate to pyruvate
Biosynthesis/Metabolism	HD1243	<i>citF</i> , citrate lyase alpha chain	TCA cycle, citrate:acetyl ACP transferase
Biosynthesis/Metabolism	HD1334	<i>sucB</i> , dihydrolipoamide succinyltransferase	TCA cycle, component of 2-oxoglutarate dehydrogenase complex
Cell Division	HD1582	<i>mukB</i> , cell division protein	necessary for chromosome partitioning and cell cycle progression
Cell Division	HD0050	<i>sufl</i> , cell division protein	periplasmic protein involved in cell division
Cell Division	HD1940	<i>xerC</i> , tyrosine recombinase	required for chromosome dimer segregation
DNA Replication/Repair	HD1731	<i>nrdA</i> , ribonucleoside-diphosphate reductase, alpha chain	catalyzes dNTP synthesis from NTPs
DNA Replication/Repair	HD0571	<i>rep</i> , ATP-dependent DNA helicase	DNA helicase involved in replication

DNA Replication/Repair	HD0626	<i>recD</i> , exodeoxyribonuclease V, alpha subunit
DNA Replication/Repair	HD1066	<i>ung</i> , uracil-DNA glycosylase
DNA Replication/Repair	HD0743	<i>mutL</i> , DNA mismatch repair protein
DNA Replication/Repair	HD1921	<i>recG</i> , ATP-dependent DNA helicase
DNA Replication/Repair	HD1121	<i>radA</i> , DNA repair protein
DNA Replication/Repair	HD2024	<i>mutS</i> , DNA mismatch repair protein
DNA Replication/Repair	HD1076	<i>recB</i> , exodeoxyribonuclease V, beta subunit
DNA Replication/Repair	HD1286	<i>uvrA</i> , excinuclease ABC subunit A
DNA Replication/Repair	HD1590	<i>deadD</i> , cold-shock DEAD box protein-A
DNA Replication/Repair	HD0973	<i>dnaB</i> , replicative DNA helicase
DNA Replication/Repair	HD1276	<i>topB2</i> , DNA topoisomerase III
Electron Transport	HD0008	<i>atpA</i> , ATP synthase alpha chain
Electron Transport	HD1393	<i>torY</i> , cytochrome c-type protein
Electron Transport	HD1158	<i>glpB</i> , anaerobic glycerol-3-phosphate dehydrogenase, subunit B
Electron Transport	HD0084	<i>lldD</i> , L-lactate dehydrogenase
Electron Transport	HD0344	<i>nrfA</i> , nitrite reductase, cytochrome c552
Electron Transport	HD1160	<i>glpC</i> , anaerobic glycerol-3-phosphate dehydrogenase, subunit C
Electron Transport	HD0074	<i>napA</i> , periplasmic nitrate reductase
Hypothetical	HD0109	hypothetical protein
Hypothetical	HD0148	conserved hypothetical protein
Hypothetical	HD0154	conserved hypothetical protein
Hypothetical	HD0155	hypothetical protein
Hypothetical	HD0256	conserved hypothetical protein
Hypothetical	HD0261	conserved hypothetical protein
Hypothetical	HD0302	conserved hypothetical protein
Hypothetical	HD0340	conserved hypothetical protein
Hypothetical	HD0521	conserved hypothetical protein
Hypothetical	HD0522	hypothetical protein
Hypothetical	HD0529	hypothetical protein
Hypothetical	HD0532	conserved hypothetical protein
Hypothetical	HD0533	hypothetical protein
Hypothetical	HD0534	hypothetical protein
Hypothetical	HD0535	hypothetical protein
Hypothetical	HD0607	conserved hypothetical protein
Hypothetical	HD0646	conserved hypothetical protein
Hypothetical	HD0921	conserved hypothetical protein, N-terminal truncated
Hypothetical	HD0922	conserved hypothetical protein
Hypothetical	HD0933	conserved hypothetical protein
Hypothetical	HD0935	conserved hypothetical protein
Hypothetical	HD0936	conserved hypothetical protein
Hypothetical	HD0966	hypothetical protein
Hypothetical	HD1061	conserved hypothetical protein
Hypothetical	HD1296	conserved hypothetical protein
Hypothetical	HD1520	conserved hypothetical protein
Hypothetical	HD1561	hypothetical protein
Hypothetical	HD1567	hypothetical protein
Hypothetical	HD1618	hypothetical protein
Hypothetical	HD1632	conserved hypothetical protein
Hypothetical	HD1655	conserved hypothetical protein
Hypothetical	HD1698	hypothetical protein
Outer Membrane Proteins	HD0045	<i>momp</i> , major outer membrane protein
Outer Membrane Proteins	HD1078	<i>ompP1</i> , outer membrane protein P1
Outer Membrane Proteins	HD1094	possible outer membrane serine protease
Regulation	HD1185	<i>relA</i> , GTP pyrophosphokinase
Regulation	HD1470	<i>cpxA</i> , sensor kinase
Ribosomal Proteins	HD1882	<i>rplA</i> , 50S ribosomal protein L1
Ribosomal Proteins	HD1968	<i>rplN</i> , 50S ribosomal protein L14

**DNA repair: endo- & exonuclease, helicase, ATPase**  
**DNA repair: excises uracil mistakenly added to DNA**  
**DNA repair: mismatch repair**  
**DNA repair: processes Holliday junction intermediates to mature form**  
 DNA repair: repair of endogenous damage  
 DNA repair: repair of mismatches in DNA  
**DNA repair: required for repair of dsDNA breaks**  
**DNA repair: UvrAB complex finds, repairs DNA lesions**  
 helix-stabilizing helicase, important for cell division  
 initiation, elongation of chromosome replication  
**topoisomerase**  
 aerobic respiration: ADP conversion from ATP with proton gradient  
 anaerobic respiration: anaerobic respiratory chain component  
 anaerobic respiration: converts G-3-P to dihydroxyacetone  
 anaerobic respiration: cytochrome synthesis  
 anaerobic respiration: electron transfer  
 anaerobic respiration: electron transfer from glycerol-3-phosphate  
 anaerobic respiration: nitrate reductase, involved in electron transfer

OmpA homolog  
 putative outer membrane protein  
 putative outer membrane protein  
 ppGpp metabolism (generates pppGpp)  
 sensor kinase of 2-component regulatory system

Ribosomal Proteins	HD1977	<i>rpsC</i> , 30S ribosomal protein S3	
Secreted Proteins	HD1156/1505	<i>lspA1</i> or <i>lspA2</i> , large supernatant protein A (conserved sequence)	antiphagocytic protein
Stress Response	HD0565	<i>clpB</i> , ATP-dependant Clp protease chain B	ATPase subunit of protease, heat shock inducible
Stress Response	HD0353	<i>htpX</i> , putative protease	probable heat shock protein
Toxins	HD0903	<i>cdtB</i> , cytolethal distending toxin protein B	part of CDT holotoxin
Transcription	HD0213	<i>srnB</i> , ATP-dependent RNA helicase protein	helicase, interacts with 23S rRNA
Transcription	HD1876	<i>rpoC</i> , RNA polymerase beta' subunit	part of RNAP catalytic core
Transcription	HD1877	<i>rpoB</i> , RNA polymerase beta subunit	part of RNAP catalytic core
Transcription	HD0604	<i>pcnB</i> , polyA polymerase	polymerase, creates 3' polyA tail, plasmid copy number protein in <i>E. coli</i>
Translation	HD0299	<i>typA</i> , GTP-binding protein, elongation factor, interacts wih ribosomes	elongation step: GTPase, mediates EPEC interactions with epithelial cells
<b>Translation</b>	<b>HD0657</b>	<b><i>fusA</i>, elongation factor G</b>	<b>elongation step: moves protein chain from A to P site of ribosome</b>
<b>Translation</b>	<b>HD0054</b>	<b><i>tufA</i>, elongation factor tu, EF-Tu</b>	<b>elongation step: promotes aa-tRNA binding to A site of ribosome</b>
Translation	HD0194	<i>rne</i> , ribonuclease E	mRNA degradation: 5S rRNA maturation, mRNA turnover
Translation	HD1588	<i>pnp</i> , polyribonucleotide nucleotidyltransferase	mRNA degradation: mRNA degradation, 3' to 5'
Translation	HD1765	<i>rumA</i> , 23S rRNA (Uracil-5-)-methyltransferase	rRNA synthesis: 23S rRNA synthesis/modification
<b>Translation</b>	<b>HD1093</b>	<b><i>cysS</i>, cysteinyl-tRNA synthetase</b>	<b>tRNA synthesis: charges tRNA with cysteine</b>
Translation	HD1942	<i>glyS</i> , glycyl-tRNA synthetase beta subunit	tRNA synthesis: charges tRNA with glycine, beta subunit of tetramer
Translation	HD1318	<i>lysS</i> , lysyl-tRNA synthetase	tRNA synthesis: charges tRNA with lysine
Transport/Uptake	HD1619	putative sodium/alanine symporter	nutrient scavanging: alanine transport
Transport/Uptake	HD0768	<i>manX</i> or <i>ptsL</i> , mannose-specific phosphotransferase IIAB component	nutrient scavanging: carbohydrate transferase in mannose-specific PTS system
Transport/Uptake	HD1824	<i>modB</i> , molybdenum ABC transporter, permease protein	nutrient scavanging: molybdenum transport
Transport/Uptake	HD1728	<i>nhaB</i> , Na <sup>+</sup> /H <sup>+</sup> antiporter protein	nutrient scavanging: sodium transport
Transport/Uptake	HD1669	<i>satA</i> , putative ABC transporter periplasmic binding protein	sialic acid transporter
Transport/Uptake	HD1670	<i>satB</i> , ABC transporter, permease protein	sialic acid transporter
Transport/Uptake	HD1528	<i>exeA</i> , type II secretory protein	Type II ( <i>sec</i> -dependent) secretion pathway component
Transport/Uptake	HD1752	<i>secD</i> , Protein-export membrane protein	Type II ( <i>sec</i> -dependent) secretion pathway component
<b>Transport/Uptake</b>	<b>HD1788</b>	<b><i>secA</i>, preprotein translocase SecA subunit</b>	<b>Type II (<i>sec</i>-dependent) secretion pathway component</b>
<b>Transport/Uptake</b>	<b>HD1886</b>	<b><i>secE</i>, preprotein translocase SecE subunit</b>	<b>Type II (<i>sec</i>-dependent) secretion pathway component</b>

\* Genes in **bold** have homologs identified during in vivo expression studies in other pathogens.

### Supplementary Data Table S3

<u>Gene ID</u>	<u>Forward Primer</u>	<u>Reverse Primer</u>
HD0192	5'-ACGGTAACAGACACTGCAGCAAAC-3'	5'-CGGCAGCTGATTCTTTCATGCCTT-3'
HD0286	5'-CCGGTACAGGGCAATATGGTAATG-3'	5'-GCTTGGGTATGTCCCTAATCTGTG-3'
HD0646	5'-AGCACGTACTAACCAAACCCGAGA-3'	5'-AGCCCTGACAAGCTAGTTCTTCCT-3'
HD0805	5'-TTAGCCACAGCAAGTGTCGGTGAT-3'	5'-TAGCCAATACCTAACCCGACCTCT-3'
HD1170	5'-GGCATTAAACGCTTGCTTCCGCATT-3'	5'-TCCACGATGACTGCTTTCCTTCCC-3'
HD1280	5'-TACGGCAACGGATATTGGTGCAGA-3'	5'-TGGCATGCCTGAAATCTGTATCACGAG-3'
HD1589	5'-AGTGGCCAGCTTCGTTTCTTAAC-3'	5'-TCAAAGTAGAGGTCTGCTCGCTCA-3'
HD1629	5'-AGAACATACCGATCCTGCTTGGCA-3'	5'-TGGTCAATTGGAAATGACACGCCT-3'
HD1655	5'-GTGACGACTACGGCGGAATATCAA-3'	5'-TGACACCACGTTTGTATTAGGTGA-3'
HD1808	5'-TTCACGCGGTGAATTAATGGCAGG-3'	5'-TCCGGCGCTTACTGAATGCTTCT-3'
HD1829	5'-GTTGAACAAGCCAAGCAAGCGACT-3'	5'-TTGCCGCTACATCTGGACGATTTG-3'