

Figure S1. Intergenic sequences between orthologous *M. PabI* homolog genes and upstream genes. Start codons for *M. PabI* homolog are in bold and stop codons for upstream genes are shaded. Under the nucleotide sequences, encoded protein sequences are shown.

(A) H-1 (*H. sp.* MIT 01-6451, 03-1614 and *H. typhlonius*)

01-6451 GCGAAAAGCTATAAAGAACTTTGCGCTTTG**TGA**ATCTGAAGCTTGAAAAATATCTATTAG-----TAGCAA
A K S Y K E L C A L *

03-1614 GCAAAAAGTTATAAAGAACTTTGCGCTTTG**TAA**ATCTGAAGCTT-AAAAATATCTATTAGCATATTTATTGCAA
A K S Y K E L C A L *

typhlonius GCAAAAAGTTATAAAGAACTTTGCGCTTTG**TAA**ATCTGAAGCTT-AAAAATATCTATTAGCATATTTATTGCAA
A K S Y K E L C A L *

01-6451 -----CATTGAGTGTAAAGTAAGAATAGAAAT**TAATG**ACAGCTCAAGCACATTTAGACAGAATCCAA
M T A Q A H L D R I Q

03-1614 TATTCAACCATTGAGAGTAAAGCGAGAACAGAA-TAA**ATG**ACAGCCCAAGCACATTTAGATAGAATCCAA
M T A Q A H L D R I Q

typhlonius TATTCAACCATTGAGAGTAAAGCGAGAACAGAA-TAA**ATG**ACAGCCCAAGCACATTTAGATAGAATCCAA
M T A Q A H L D R I Q

(B) H-2 (*H. muridarum* ST1) & C-4 (*C. concisus* UNSW2)

ST1 GATTCTATCTTTAAGCGGGTGTTTGGGGAGGGAGAT**GA**ATAAAAAAACATTTAGTTTTGGAATC//
D S I F K R V F G E G R *

ST1 //AGACAATGTTTAAAGGGTTTTAGATA--**ATG**ACAGCACAAAACACCTAAGCACCAAGGAA
M T A Q T H L S T K E

UNSW2 GATAAAATTTTAAAGAGATATTTGAAAA**TAATG**TCGATACAAAGACACTTAAATAAAATCGAT
D K I F K R V F E K *

M S I Q R H L N K I D

(C) H-3 (*H. bilis* ATCC 43879, *H. cinaedi* CCUG 18818, ATCC BAA-849, PAGU611, *C. upsaliensis* RM3195) & C-1 (*C. sp.* MIT 97-5078)

ATCC 43879 TCAAAACCCCCGGGAACATTGAGTGGGAGTAGAGATAATTC**ATG**ACAAAACCAACGCATTTAAGCACACAAGAA
S K P P G T I E W E * M T K P T H L S T Q E

CCUG 18818 TCAAAACCCCCGGGAACGATTGAGTGGGAGTAGAGATAATTC**ATG**ACAAAACCAACGCATTTAAGCACACAAGAA
S K P P G T I E W E * M T K P T H L S T Q E

BAA-849 TCAAAACCCCCGGGAACGATTGAGTGGGAGTAGAGATAATTC**ATG**ACAAAACCAACGCATTTAAGCACACAAGAA
S K P P G T I E W E * M T K P T H L S T Q E

PAGU611 TCAAAACCCCCGGGAACGATTGAGTGGGAGTAGAGATAATTC**ATG**ACAAAACCAACGCATTTAAGCACACAAGAA
S K P P G T I E W E * M T K P T H L S T Q E

RM3195 TCAAAACCCCCGGGAACGATTGAGTGGGAGTAGAGATAATTC**ATG**ACAAAACCAACGCATTTAAGCACACAAGAA
S K P P G T I E W E * M T K P T H L S T Q E

MIT 97-5078 AGCAAACCCCCAGGAACCATTTGAGTGGGAGTAGGGA-----**ATG**ACGATACAAGCACATTTAAATAAAATCCAA
S K P P G T I E W E * M T I Q A H L N K I Q

(D) H-4 (*H. suis* HS1 and *H. heilmannii* ASB1.4)

HS1 CCTTTGAAGTGATTTTACAGACTATGCAAGAGGTGTTGTT**ATA**AAACCTTCACATCTGCATTACTCCGAGCAATCAATTT
P F E V I L Q T M Q E V L L *
M K P S H L H Y S E Q I N L

HS1 CCTTTGAAGTGATTTTACAGACTATGCAAGAGGTGTTGTT**ATA**AAACCTTCACATCTGCATTACTCCGAGCAATCAATTT
P F E V I L Q T M Q E V L L *
M K P S H L H Y S E Q I N L

ASB1.4 ACAGCTCCTTAGTGAAGCAACCATGCAAGAAGTCTAA**TTGA**-----GCCACCTAAGCCCCCTTGAAAAGATCAATCT
N S S L V Q A T M Q E V L N *
M S H L S P L E K I N L

(E) H-9 (*H. pylori* F16, J99, *H. acinonychis* Sheeba, *H. cetorum* 00-7128 and 99-5656)

F16 CATGCTAAAGATGTGAAGCAGCTAGATGTG**TAA**GGGATT**TTG**ATTA---ACCCTAACATTTAAATAAGCGAGAG
H A K D V K Q L D V * M I N P K H L N K R E

J99 CATGCCAAAGATGTGAAGCAACTAGATGTG**TAA**GGGTT**TTG**ATTAGAACCCCAAGCATTTAACCAAGCAAGAG
H A K D V K Q L D V * M I R T P K H L T K Q E

Sheeba CATGCTAAAGATGTGAACAATTGGATGTG**TAA**GGGATTT**ATC**ACTACCCCTAACATTTAACCAAGCAAGAG
H A K D V K Q L D V * M T T P K H L T K Q E

00-7128 CATGCTAAAGATGTGAACAACCTAGATGTG**TAA**GGAAAATTT**ATC**ACAACCCCTAACATTTAACCAACAAGAG
H A K D V K Q L D V * M T T P K H L T K Q E

99-5656 CATGCTAAAGATGTGAACAACCTAGATGTG**TAA**ATAATGTTTTGTACCGGCAATTTAGTTCGGTATGGGTATCT//
H A K D V K Q L D V *
CCCTAGAGTGAGCTATTTTTATTACTAAGGAGTTTTG**ATC**ACAACCCCTAACATTTAACCAACAAGAA
M T T P K H L T K Q E

(F) C-3 (*C. jejuni* ATCC 33560, *coli* LMG 23336, *coli* H56)

33560 ATCAAGATTTGCAAGAATATTTAAGGACTAGTATT**GATG**CAAAACCATCTTAACAGAGAAATAATGATT
I K S L Q E Y L K D * M Q N H L N R E I M I

LMG23336 ATCAAGATTTGCAAGAATATTTAAGGACTAGTATT**GATG**CAAAACCATCTTAACAGAGAAATAATGATT
I K S L Q E Y L K D * M Q N H L N R E I M I

H56 ATCAAGATTTGCAAGAATATTTAAGGACTAGTATT**GATG**CAAAACCATCTTAACAGAGAAATAATGATT
I K S L Q E Y L K D * M Q N H L N R E I M I