Supporting Information

Chiu *et al.* 10.1073/pnas.0805968105

SI Text

Cardiovirus Screening of Stool from the SIFT Cohort. From January 2000 to July 2004, patients (mostly children) with diarrhea, vomiting, or both were seen at 1 of 15 community clinics in northern California. Eligible households were scheduled for two home visits. During the first visit, the patient and consenting household members were interviewed regarding symptoms (nausea, vomiting, or unspecified GI symptoms), onset, and duration of gastroenteritis within the previous 10 days. During the second visit 3 to 5 months after the first visit, the patient and consenting household members were interviewed regarding resolution of GI and residual symptoms (if any). No information was obtained on fever or other constitutional symptoms. During each visit, biological specimens (serum, stool, or both) were obtained from the affected patient and consenting household members for diagnostic testing. Stools were preferentially collected over sera in children less than 2 years old, because *H. pylori* serodiagnostics were considered less accurate in this group. In older individuals, stools were obtained only if participants volunteered both stool and serum specimens or refused blood draw. Specimens were collected under protocols approved by the Stanford Institutional Review Board. Subjects participating in the last few years of the study were asked to consent to unspecified research using their de-identified biological specimens. In total, 4333 individuals in the SIFT study provided biological specimens; 3063 (71%) of these participants provided consent for use of de-identified specimens. Among the 3063 subjects who consented to further use of biological specimens, 774 stool specimens were obtained from 514 individuals; of those, 751 specimens from 498 subjects were available for study. Those who provided consent for further use of specimens were younger (mean: 19.5 vs. 21.1 years, respectively, $P = 0.004$) and more likely to be Hispanic (86% vs. 83%, respectively, $P = 0.002$) than those who did not provide consent. Those who consented to unspecified use of specimens and those who did not were equally likely to provide stool specimens (17 vs. 18%, respectively, $P = 0.3$).

Among the 3063 subjects who consented to further use of biological specimens, 774 stool specimens were obtained from 514 individuals; of those, 751 specimens from 498 subjects were available for study. Not unexpectedly based on the study design, those who provided stool samples were far younger (mean of 3.2 years) than those who did not (mean of 22.7 years, $P < 0.001$); 80% of those who provided stool samples were children <2 years of age. Compared with those who provided only serum, those who provided a stool sample were also more likely to be male (52% vs. 40%, $P < 0.001$). Since young children were more likely to have GI symptoms than older subjects [Perry S, de la Luz Sanchez M, Hurst PK, Parsonnet J (2005) Household transmission of gastroenteritis. *Emerg Infect Dis* 11:1093–1096], diarrhea and vomiting were far more common among those who provided stool samples at first visit (11% vomiting, 19% diarrhea, and 56% both vomiting and diarrhea) than among those who did not (8% vomiting, 11% diarrhea, and 15% both vomiting and diarrhea; odds ratio for any symptoms = 11.7; 95% confidence interval: 8.8, 15.6). This difference was less marked when adjusted for age group (<2 years, 2–17 years, 18+ years), although those individuals who provided stool were still more likely to have GI symptoms than those who did not (summary odds ratio = 2.4, 95% confidence interval: 1.5, 3.7).

A L protein

```
HTCV-UC1   MA-------CKHGYP-LMCPLCTALDKTSDGLFTLLFDNEWYPTDLLTVDLEDEVFYPDD----PH-MEWTDLPLIQDIEMEPQ  71
HTCV-Saf   MA-------CKHGYP-FLCPLCTAIDISADGSFALLFDNEWYPTDLLTVDLDDDVFHPPD----CV-MEWTDLPLIQDVLMEPQ  71
Vilyuisk   MA-------CKHGYP-DVCPICTAIDVTPGFEYLLLADGEWFPTDLLCVDLDDDVFWPSDSSNQSQTMEWTDIPLICDTVMEPQ  76
TMEV-DA    MA-------CKHGYP-DVCPICTAVDVTPGFEYLLLADGEWFPTDLLCVDLDDDVFWPSNSSNQSETMEWTDLPLVRDIVMEPQ  76
NGS910     MA-------CIHGYP-SVCPICTAIDKSSDGMYLLLADNEWFPADLLTMDLDDDVFWPNDESDVSETMDWTDLPFILDTIMEPQ  76
EMCV       MATTMEQEICAHSMTFEECPKCSALQYRNGF-YLLKYDEEWYPEESLT-DGEDDVFDP-------------DLDM--EVVFETQ  67
```

Zinc finger — Acidic domain — Ser/Thr rich domain

L* protein

```
HTCV-UC1   TDIRLCALFALLSTKLRTDFSPFCSTMNGTQLTYX
HTCV-Saf   TDIRFCALFALLLTSLQMDLLLYYLTMNGTRLTSLLLTWTTTCFIPRIV------WNGLIYHX
Vilyuisk   MDTQTCALFAQPLTLLPALNICSWRTENGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPRHLTRVTPS
TMEV-DA    MDTQMCALFAQPLTLLPDLNICSWQTVNGSQRTFFVWTWTMTSSGLRTRAINLKQWNGLTYRSYAILSWNPRETPLHLTRVTPS
NGS910     TDTQTCALFAQPLTLLPTLNICSWQTENGSLRTFFVWTWTMTSSGLRTRALNLKQWNGLMYRSYAILSWNPREMPRHLIRVTPS
EMCV       MDTQACVLFAQPLTKVPTECICSWQITNGSQRIFLLWTWMMTSSGLMTRAMCLRQWTGLTFRSYSILSWNPRETPRHLTRVTPS
```

B CD loops

```
                    I                    II
HTCV-UC1   LTPLPSNRLDDS-----TYGLAEQHRWLSFP---TDTKQTPPYKTKQD  111
HTCV-Saf   LTPLPSDRLKEN-----EFGLDEQHRWLSFQ---SATSSTPPYRTKQD  111
Vilyuisk   LTPLPSYSPDRPGQSPDTSKAPIQWRWISAVTESGTVSNTFPTRTRQD  118
TMEV-DA    LTPLPSYCPDSS-SGPVRTKAPVQWRWVRSG---GANGANFPLMTKQD  113
NGS910     LTPLPSYAPDST-TGPTETQAPIQWRWLRGT---SDGSTTFPLMTKQD  113
EMC        LTPGPQFDPAYD----QLRPQRLTEIWGNGN---EETSKVFPLKSKQD  111
```

EF loops

```
                I                          II
HTCV-UC1   PEFDTSPYNATTEPTKAVPFQMDTQWQS--GKLLGHSYESTTLQGLRPLALNHQN  184
HTCV-Saf   PEFDTSSYSAVDDPIGEEPFKVDTTWQT--GSLRGHSYEDKSTQTLRPLALNHQN  184
Vilyuisk   PEFYTGTGVATSGQEPNKVFLMDTTWQEPQAAPTGFRYDGKN----GFFTLNHQN  182
TMEV-DA    PEFYTGKGTKSGTMEPSDPFTMDTTWRSPQSAPTGYRYDRQA----GFFAMNHQN  182
NGS910     PEFYTGHTPVTGTTEPQTPFTMDSSWQTPQQNPVGFRYDGRT----GYFALNHQN  182
EMC        PEYPT-----------LDAFAMDNRWSK-DNLPNGTRTQTNKK---GPFAMDHQN  171
```

Fig. S1. Alignment of cardiovirus proteins. (*A*) L protein and L* proteins. The zinc-finger domain of the L protein is highlighted in pink, the acidic domain in yellow, and the Ser/Thr rich domain in cyan. Fully conserved residues of the L* protein are highlighted in green, and stop codons are designated with an X. For Vilyuisk virus, TMEV-DA, Theiler-like NGS910 virus, and EMCV, only the first 84 aa of the L* protein (≈156 aa) are shown. (*B*) CD and EF loops. The regions corresponding to the CD loops of VP1 (I and II) and EF loops of VP2 (I and II) are highlighted in light tan. Sialic acid binding residues in the TMEV VP2 protein are shown boldfaced in red.

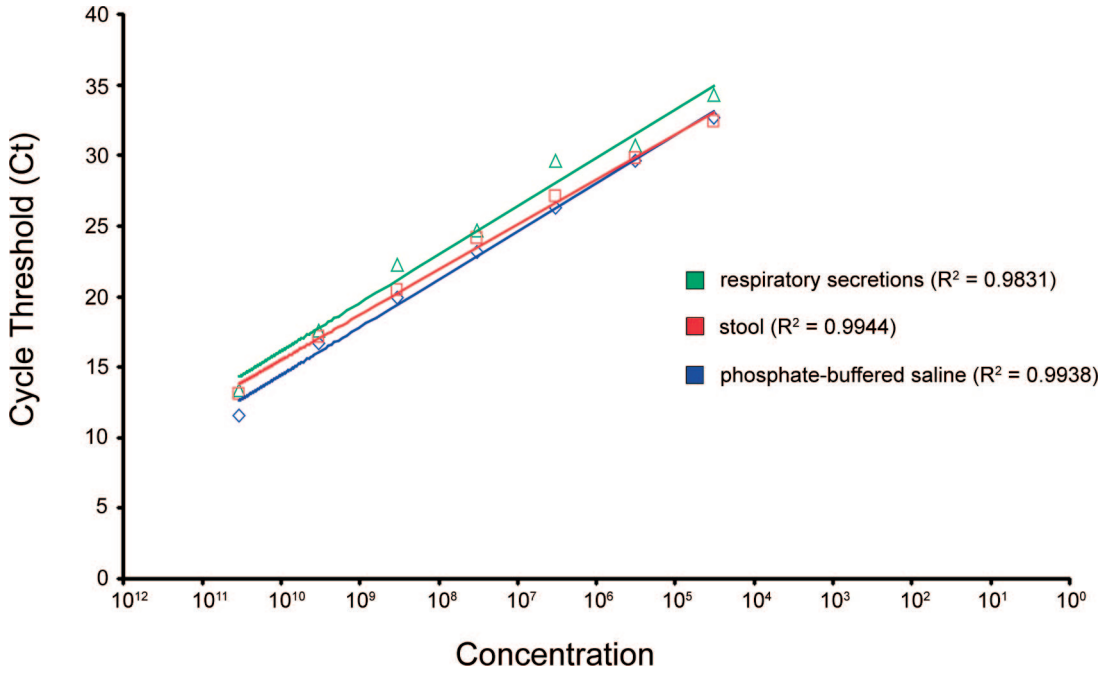

Fig. S2. Log-log plots of standard curves generated using a real-time RT-PCR assay for cardioviruses on spiked specimens. Specimens were prepared by extraction of RNA from seven serial dilutions of *in vitro* transcribed UC1 mRNA spiked into pooled respiratory secretions (green), pooled stool suspensions (red), and phosphate-buffered saline (blue). Each pool consists of 10 randomly selected cardiovirus-negative specimens, and each data point is an average of three independent replicates.

Table S1. Primers used for cardiovirus screening and VP1 sequencing by RT-PCR

Primer	Sequence	UC1 nucleotide position
CardioUTR-1F	5′-CATTTTCCGGCCCAGGCTAAG-3′	121-141
CardioUTR-2R	5′-GTGAACAAGCGGCAAGGGAG-3′	222-203
CardioUTR-3R	5′-GCTCACAGCAGTGGATCTTATCC-3′	729-707
CardioVP1-1F	5′-ATGCAAGCCACCTATGCTATTTGGG-3′	2615-2639
CardioVP1-2F	5′-GGTGGGGATGACTTTACCCTCAGAATGCC-3′	2825-2853
CardioVP1-3R	5′-TTGTAAGTGAATTGAATGATTTCATCTG-3′	3771-3744
CardioVP1-4R	5′-TATTGACAAACTGTTCTGCCATG-3′	3798-3776