AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

GENETIC DIVERSITY AND EVOLUTION

Novel Synthesis and Phenotypic Analysis of Mutant Clouds for Hepatitis E Virus Genotype 1

Shubhra Agarwal,[a] Prasith Baccam,[b] Rakesh Aggarwal,[c] Naga Suresh Veerapu[a]

[a]Virology Section, Department of Life Sciences, Shiv Nadar University, Gautam Buddha Nagar, Uttar Pradesh, India

[b]IEM, Inc., Bel Air, Maryland, USA

[c]Gastroenterology, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, Uttar Pradesh, India

ABSTRACT Many RNA viruses exist as an ensemble of genetically diverse, replicating populations known as a mutant cloud. The genetic diversity (cloud size) and composition of this mutant cloud may influence several important phenotypic features of the virus, including its replication capacity. We applied a straightforward, bacterium-free approach using error-prone PCR coupled with reverse genetics to generate infectious mutant RNA clouds with various levels of genetic diversity from a genotype 1 strain of hepatitis E virus (HEV). Cloning and sequencing of a genomic fragment encompassing 70% of open reading frame 1 (*ORF1*) or of the full genome from variants in the resultant clouds showed the occurrence of nucleotide mutations at a frequency on the order of 10^{-3} per nucleotide copied and the existence of marked genetic diversity, with a high normalized Shannon entropy value. The mutant clouds showed transient replication in cell culture, while wild-type HEV did not. Cross-sectional data from these cell cultures supported the existence of differential effects of clouds of various sizes and compositions on phenotypic characteristics, such as the replication level of (+)-RNA progeny, the amounts of double-stranded RNA (a surrogate for the rate of viral replication) and ORF1 protein, and the expression of interferon-stimulated genes. Since mutant cloud size and composition influenced the viral phenotypic properties, a better understanding of this relationship may help to provide further insights into virus evolution and prediction of emerging viral diseases.

IMPORTANCE Several biological or practical limitations currently prevent the study of phenotypic behavior of a mutant cloud *in vitro*. We developed a simple and rapid method for synthesizing mutant clouds of hepatitis E virus (HEV), a single-stranded (+)-RNA [ss(+) RNA] virus, with various and controllable levels of genetic diversity, which could then be used in a cell culture system to study the effects of cloud size and composition on viral phenotype. In a cross-sectional analysis, we demonstrated that a particular mutant cloud which had an extremely high genetic diversity had a replication rate exceeding that of wild-type HEV. This method should thus provide a useful model for understanding the phenotypic behavior of ss(+) RNA viruses.

KEYWORDS phenotype, error-prone PCR, genetic diversity, hepatitis E virus, mutant cloud, quasispecies, random mutagenesis, reverse genetics

Genomes of positive-sense, single-stranded [ss(+)] RNA viruses have a unique dual role, i.e., to serve as the template for both translation and replication. Once such a virus enters the host cell, the ss(+) viral genome reaches the cytoplasm and is translated into structural and nonstructural (NS) proteins. The NS proteins, especially the RNA-dependent RNA polymerase (RdRp), and host factors then generate nascent ss(+) RNA genomes via the synthesis of negative-strand RNA, a replication intermediate

Received 7 November 2017 **Accepted** 7 November 2017

Accepted manuscript posted online 22 November 2017

Citation Agarwal S, Baccam P, Aggarwal R, Veerapu NS. 2018. Novel synthesis and phenotypic analysis of mutant clouds for hepatitis E virus genotype 1. J Virol 92:e01932-17. https://doi.org/10.1128/JVI.01932-17.

Editor Julie K. Pfeiffer, University of Texas Southwestern Medical Center

Address correspondence to Naga Suresh Veerapu, nagasuresh.veerapu@snu.edu.in.

(1). This replication mechanism of ss(+) RNA viruses is inherently prone to errors, partly due to the lack of a proofreading ability in their RdRps (2), with estimated mutation rates of 10^{-3} to 10^{-5} substitution per nucleotide per genome replication (3). This results in the creation of a set of genetically distinct, albeit closely related, genetic variants which are capable of self-replication, a phenomenon termed variously "quasispecies," "mutant swarms," or "mutant clouds" (4–9).

The complexity and composition of the mutant clouds generated *in vivo* have been proposed to contribute to the adaptability, persistence, and tissue tropism of these viruses, as well as to antiviral drug resistance (10). Recent years have seen a growing interest in studying mutant clouds to understand the biology of RNA viruses as well as their role in the pathogenesis of viral infection and disease (11, 12). The available data suggest that mutant clouds are dynamic virus populations that have evolved to attain mutation rates that are optimally suited to the particular environment in which the replication is occurring. These clouds are believed to be influenced by intrapopulation interactions and host selection pressures and to constantly evolve to adapt to the ever-changing environment. However, the transient equilibria achieved *in vivo* are difficult to observe (13, 14).

Reverse genetics approaches have helped to advance our understanding of the effects of genetic variations in ss(+) RNA viruses on viral replication, pathogenesis, drug resistance, and vaccine development (15). In the classical approach, full-length wild-type (WT) or mutant virus cDNA clones with one or a few mutations of interest are constructed and used as templates for the synthesis of infectious ss(+) RNA (16–18). Given the recent knowledge about mutant clouds, it seems important to study the effects of full-length genome diversity, i.e., the cocirculation of several slightly different viral mutants at various relative concentrations. However, this has been hampered by the limited availability of tools for generating such clouds.

Error-prone PCR (Ep-PCR) is a powerful technique for *in vitro* generation of genetic diversity. It has been used widely to implement directed evolution and has shown remarkable success in protein engineering (reviewed in reference 19). It uses the tendency of low-fidelity *Taq* DNA polymerases to incorporate incorrect nucleotides during nucleotide chain extension in a nontargeted manner. The genetic diversity so produced can be increased further by varying the magnesium concentration, supplementation with manganese ions, and use of unbalanced amounts of the four deoxynucleotide triphosphates (dNTPs), a small template amount, and/or a large number of amplification cycles (20). Our literature search revealed that the longest sequence which had been diversified using Ep-PCR was ~2.5 kb long (21). This limits our ability to mutagenize the full-length genome of even the smallest human ss(+) RNA virus.

Hepatitis E is a major public health concern in developing countries and is considered a reemerging disease in developed countries. Its causative agent, hepatitis E virus (HEV), has at least the following two transmission profiles: (i) in developing countries, human-to-human transmission of predominantly genotype 1 (and, less frequently, genotype 2) HEV via fecal contamination of drinking water; and (ii) in developed countries, zoonotic transmission from an animal reservoir of genotype 3 (or 4) HEV via consumption of undercooked contaminated meat (22). HEV is a nonenveloped virus with an approximately 7.2-kb ss(+) RNA genome. It consists of a short 5′ untranslated region (UTR) and a 3′ UTR which terminates with a poly(A) tail and contains three partially overlapping open reading frames, ORF1 to ORF3, encoding a putative NS polyprotein, the viral capsid protein, and a small phosphoprotein, respectively (23). Genotype 1 HEV is extremely difficult to propagate *in vitro*. HEV appears to follow an error-prone replication process, naturally circulating as a heterogeneous population in human hosts (24–27). However, information on error rates and the natural evolutionary dynamics of HEV is limited, partly due to the short duration of viremia in most patients (28).

In the current study, we used a rapid and versatile approach, creation of a full-length mutant RNA cloud (FL-MRC), for *in vitro* generation of infectious mutant clouds of full-length HEV RNA. In brief, we used Ep-PCR to generate mixtures of genetically

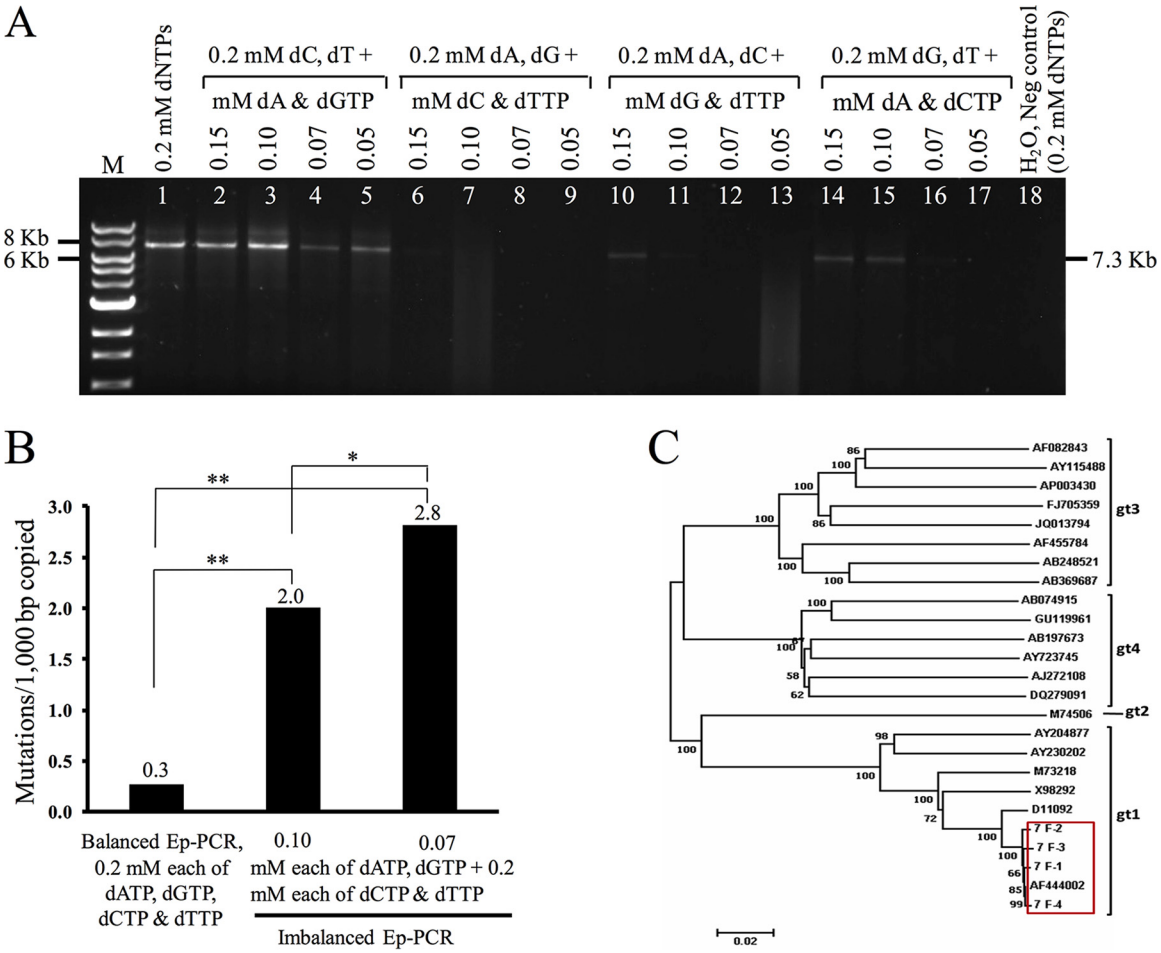

FIG 1 Agarose gel electrophoresis profiles of Ep-PCR products and genetic diversity of HEV amplicons. (A) Agarose gel analysis of Ep-PCR products. The HEV genome was amplified from the pSK-E2 plasmid by use of either balanced dNTP concentrations (lane 1) or combinations of dinucleotides at unbalanced concentrations (lanes 2 to 17), as indicated. Detailed Ep-PCR conditions are given in Materials and Methods. The expected size of amplicons was 7.2 kb. The products were separated in a 0.8% agarose gel and visualized by staining with ethidium bromide. Lane M, 1,000-bp DNA size ladder (sizes given in kilobase pairs). (B) Proportions of mutations per 1,000 bp in full-length amplicons arising from Ep-PCRs as determined by DNA sequencing. Statistically significant differences are indicated by asterisks (**, $P < 0.001$; *, $P < 0.05$ [Fisher's exact test]). (C) Phylogenetic analysis of full-length HEV genomes. The neighbor-joining tree shows reference full-length HEV genome sequences for genotypes 1 to 4 and full-length genome sequences from this study. Support values for branches supported by at least 60% of bootstrap replicates are shown next to the nodes. Full-length genome sequences (gt1 to gt4) and the wild-type Sar55 sequence (accession no. AF444002) are indicated in a red box. The scale bar represents the genetic distance.

diverse full-length HEV genomes, producing infectious ss(+) RNA mutant clouds of various sizes in sufficient quantities to undertake phenotypic analysis. When it was expressed in hepatoma cells, a highly complex, diverse mutant HEV RNA cloud showed remarkable phenotypic differences in terms of viral RNA replication, protein accumulation, and induction of interferon-stimulated genes (ISGs).

RESULTS

Virus FL amplicons obtained through Ep-PCR and sequences diversified with an unbalanced dNTP pool. Misincorporation of a nucleotide during PCR amplification can be followed either by further extension of the erroneous strand to a complete amplicon or by falling of the erroneous incomplete strand from the template, reducing the efficiency of the PCR and the amount of amplicon generated. Therefore, we first assessed the efficiency of Ep-PCR for amplification of full-length (FL) amplicons. Balanced PCR for amplification of HEV Sar55 FL amplicons had an average yield of 3.9 ± 0.2 ng/μl, with a PCR efficiency of 0.164 (Fig. 1A, lane 2, and Table 1). Imbalanced Ep-PCRs (with two dNTPs at 0.20 mM [each] and the other two dNTPs at 0.15, 0.10, 0.07,

TABLE 1 Error-prone PCR product yields and efficiencies and numbers of variants estimated by PEDEL-AA analysis under various mutagenesis reaction conditions

Ep-PCR conditions	Avg Ep-PCR product concn (ng/μl) ± SD	Avg Ep-PCR efficiency ± SD	PEDEL-AA result[a]			
			Library size	No. of variants with no indels or stops	No. of unique protein variants	% unique protein variants
Balanced (0.20 mM [each] dATP, dGTP, dCTP, and dTTP)	3.9 ± 0.20	0.164 ± 0.0020	5.00×10^6	4.44×10^6	3.05×10^5	6.8
Imbalanced (0.20 mM [each] dCTP and dTTP and indicated concn [mM] [each] dATP and dGTP)						
0.15	4.1 ± 0.18	0.166 ± 0.0010	ND	ND	ND	ND
0.10	4.3 ± 0.10	0.169 ± 0.0008	5.45×10^6	4.82×10^6	4.11×10^6	85
0.07	2.9 ± 0.20	0.153 ± 0.0027	4.45×10^6	2.20×10^6	2.10×10^6	95
0.05	3.1 ± 0.15	0.156 ± 0.0020	ND	ND	ND	ND

[a]A nucleotide mutation matrix, Poisson and PCR distributions for 30 cycles, and Ep-PCR efficiency were input for PDEDEL-AA analysis (http://guinevere.otago.ac.nz/stats.html). The sequence analyzed was the sequence of nucleotides 26 to 3,286 (total of 3,261 nucleotides; corresponds to amino acids 1 to 1,087 of the ORF1 protein). Nucleotide numbering is based on that of the wild-type Sar55 genome sequence (GenBank accession number AF444002). ND, not determined.

or 0.05 mM [each]) were also successful in some cases, although the PCR product yields differed. More specifically, amplification occurred with 0.2 mM (each) dCTP and dTTP and 0.15, 0.10, 0.07, or 0.05 mM (each) dATP and dGTP (Fig. 1A, lanes 2 to 5), but not when the concentration of dATP and dGTP was reduced to 0.025 mM (each) (data not shown). Table 1 shows the average yields and Ep-PCR efficiencies of various imbalanced Ep-PCRs. Imbalanced Ep-PCRs using 0.20 mM (each) dCTP and dTTP (Fig. 1A, lanes 6 to 9), dATP and dCTP (lanes 10 to 13), or dGTP and dTTP (lanes 14 to 17), with various lower concentrations of the other two dNTPs, had lower or undetectable yields and PCR efficiencies.

The 3.5-kb fragments derived after XhoI and XbaI digestion of FL amplicons obtained by balanced Ep-PCR and imbalanced Ep-PCRs (0.2 mM [each] dCTP and dTTP and 0.10 mM [each] dATP and dGTP [AG0.10-CT0.20] and AG0.07-CT0.20) were cloned into the pSK-E2 vector at compatible restriction sites. In addition, the FL amplicon obtained by the AG0.07-CT0.20 imbalanced Ep-PCR was cloned into a T-vector, and individual clones were sequenced. Table 2 shows the numbers of clones and nucleotides analyzed. The amplicons from balanced Ep-PCR had 0.3 mutation/kb (8 mutations in 29,808 bases); in comparison, the mutation frequencies in imbalanced Ep-PCR amplicons were significantly higher, at 2.0 mutations/kb (63 mutations in 31,338 bases) for AG0.10-

TABLE 2 Mutant cloud analysis by molecular cloning and Sanger sequencing of full-length amplicons of HEV amplified by Ep-PCR

Ep-PCR conditions	No. of clones analyzed	No. of nucleotides analyzed	No. of mutations[a] observed	No. of transitions/no. of transversions (ratio)	No. of nonsynonymous mutations/no. of synonymous mutations (ratio)
Balanced (0.20 mM [each] dATP, dGTP, dCTP, and dTTP)	9	29,808	8	7/1 (7)	4/4 (1.0)
Imbalanced (0.20 mM [each] dCTP and dTTP and indicated concn [mM] [each] dATP and dGTP)					
0.10	9	31,338	63	49/14 (3.5)	44/19 (2.3)[e]
0.07[b]	11	51,848	146[c]	90/52 (1.7)[d]	95/47 (2.0)[e]
Total (avg)				146/67 (2.18)	143/70 (2.04)

[a]The number of total mutations was determined relative to the wild-type Sar55 genome sequence (GenBank accession number AF444002) for sequences of the corresponding amplicons.
[b]Includes sequence data for four full-length genomes.
[c]Includes one insertion in one full-length genome and three deletions in one full-length genome.
[d]Excludes one nucleotide insertion and three deletions.
[e]Statistically significant ($P < 0.001$) for number of nonsynonymous substitutions relative to that for balanced Ep-PCR by Fisher's exact test.

TABLE 3 Mutation distribution among HEV clones obtained by balanced and imbalanced error-prone PCRs as analyzed by Sanger sequencing

Ep-PCR conditions	Clone ID	No. of nt analyzed	No. of mutations	No. of transitions/ no. of transversions	No. of nonsynonymous mutations
Balanced (0.20 mM [each] dATP, dGTP, dCTP, and dTTP)	2_C1	3,312	0		
	2_C2	3,312	1	1/0	1
	2_C3	3,312	2	2/0	0
	2_C4	3,312	0		
	2_C5	3,312	1	1/0	1
	2_C6	3,312	2	1/1	0
	2_C7	3,312	1	1/0	1
	2_C8	3,312	1	1/0	1
	2_C9	3,312	0		
Total	9	29,808	8	7/1	4
Imbalanced (0.10 mM [each] dATP and dGTP and 0.20 mM [each] dCTP and dTTP)	1_C1	3,482	3	3/0	2
	1_C2	3,482	4	4/0	3
	1_C4	3,482	9	7/2	5
	1_C5	3,482	7	3/4	7
	1_C6	3,482	5	5/0	3
	1_C7	3,482	10	10/0	7
	1_C8	3,482	6	3/3	5
	1_C9	3,482	9	6/3	6
	1_C10	3,482	10	8/2	6
Total	9	31,338	63	49/14	44
Imbalanced (0.07 mM [each] dATP and dGTP and 0.20 mM [each] dCTP and dTTP)	7_F-1	7,208	17[a]	8/8	13
	7_F-2	7,208	22	14/8	12
	7_F-3	7,208	26	16/10	17
	7_F-4	7,208	15[b]	9/3	7
	7_C4	3,288	8	6/2	7
	7_C6	3,288	7	4/3	5
	7_C8	3,288	12	7/5	8
	7_C9	3,288	11	6/5	6
	7_C10	3,288	8	5/3	7
	7_C11	3,288	10	7/3	7
	7_C13	3,288	10	8/2	6
Total	11	51,848	146	90/52	95

[a]Includes one insertion.
[b]Includes three deletions.

CT0.20 PCR and 2.8 mutations/kb (146 mutations in 51,848 bases) for AG0.07-CT0.20 PCR ($P < 0.001$ for both compared to balanced Ep-PCR and $P < 0.05$ [Fisher's exact test] for comparison between these two imbalanced Ep-PCRs) (Fig. 1B).

Mutations were biased toward transitions and nonsynonymous mutations. Under balanced Ep-PCR conditions, the ratio of number of transitions to number of transversions in the gene region sequenced was 7.0 (Tables 2 and 3). For imbalanced Ep-PCRs, this ratio was lower, i.e., 3.5 and 1.7 for AG0.10-CT0.20 and AG0.07-CT0.20 amplicons, respectively (Tables 2 and 3). The two imbalanced Ep-PCRs also had higher ratios of nonsynonymous to synonymous mutations (*dN*/*dS*), i.e., 2.0 and 2.3, respectively, than that of the balanced Ep-PCR (1.0). All 8 mutations (100%) in the amplicons from the balanced Ep-PCR and all 63 mutations (100%) from the AG0.10-CT0.20 imbalanced Ep-PCR were substitutions, and none of these led to premature stop codons, suggesting that a large proportion of these substitutions could result in viable mutant proteins. In contrast, the 146 mutations in the AG0.07-CT0.20 amplicon included one insertion (3′ UTR of clone 7_F-1) and three deletions (coding region of clone 7_F4), suggesting that a proportion of mutants in this amplicon did not code for viable proteins (see Fig. S1 in the supplemental material). Fifty percent of substitutions (4 of 8 substitutions) induced under balanced Ep-PCR conditions were nonsynonymous (Table 2). In contrast, a large proportion of substitutions induced under AG0.10-CT0.20

TABLE 4 Mutant cloud complexity indices, including mutation frequencies, Hamming distances, and normalized Shannon entropy values

Ep-PCR conditions	Mutant cloud complexity index		
	Mutation frequency[a]	Mean Hamming distance[b,c]	S_n value[c,d]
Balanced (0.20 mM [each] dATP, dGTP, dCTP, and dTTP)	2.68×10^{-4}	0.88	0.48
Imbalanced (0.20 mM [each] dCTP and dTTP and indicated concn [mM] [each] of dATP and dGTP)			
0.10	2.01×10^{-3}	6.66	1.00
0.07	2.81×10^{-3}	8.36	1.00

[a]Mutation frequency is the total number of mutations found divided by the total number of nucleotides sequenced and includes the complete sequence data given in Table 1.

[b]The mean Hamming distance was calculated by averaging the numbers of nucleotide differences in individual clones relative to the wild-type sequence.

[c]Sequences analyzed included nucleotides 26 to 3,286 (total of 3,261 nucleotides of ORF1). Nucleotide numbering is based on that of the wild-type Sar55 genome sequence (GenBank accession number AF444002).

[d]The normalized Shannon entropy (S_n) value was calculated by using the following formula: $S_n = -\{\sum_i [p_i \ln(p_i)]\}/(\ln N)$, where P_i is the frequency of individual clones and N is the number of clones sequenced. S_n values can range from 0 (completely homogeneous) to 1 (completely heterogeneous).

(40 of 63 mutations [70%]) or AG0.07-CT0.20 (95 of 142 mutations [67%]) conditions were nonsynonymous.

Mutation frequency and Shannon entropy provided a snapshot of genetic diversity in FL sequences. A nearly 1-log increase in mutation frequency, from 2.68×10^{-4} per nucleotide copied (0.3 mutation/kb) in balanced Ep-PCRs to 2.81×10^{-3} per nucleotide copied (2.8 mutations/kb) in highly imbalanced Ep-PCRs, suggests that the low fidelity attributed to *Taq* DNA polymerase is dNTP concentration dependent (Tables 2 and 3). The normalized Shannon entropy (S_n) was calculated from the frequencies of nucleotide substitutions measured across the gene region. In imbalanced Ep-PCRs, the S_n measured 1.0, indicating maximum viral diversity (Table 4).

Phylogenetic evaluation of genetically diversified viral FL sequences. We sought to analyze whether our Ep-PCR had perturbed the sequence space enough to alter virus phylogenetic grouping. Anticipating a high genetic diversity in AG0.07-CT0.20 amplicons, we cloned and sequenced four full genomes. All four sequences were found to be genetically close to the parental strain (Sar55), within the genotype 1 clade (Fig. 1C).

PEDEL-AA predicted that imbalanced Ep-PCR generated highly diverse mutant libraries. Analysis by use of the Programme for Estimating Diversity in Error-Prone PCR Libraries—Amino Acid (PEDEL-AA) (Table 1) showed that library sizes were related to respective amplicon yields and PCR efficiencies of various Ep-PCRs. Unique protein variants with no insertions, deletions, or stops were predicted to comprise 6.8% of the balanced Ep-PCR library and 85% and 95% of the AG0.10-CT0.20 and AG0.07-CT0.20 imbalanced Ep-PCR libraries, respectively, indicating that the imbalanced Ep-PCRs generated highly diverse protein-encoding libraries. Our sequence data and PEDEL analyses showed that the majority of sequences represented unique variants.

PAQ applied to clonal analysis of mutant libraries. On examination of the data from clones by use of partition analysis of quasispecies (PAQ), no distinct clusters were identified among the sequences within the clouds. The clones from balanced Ep-PCR (AG0.20-CT0.20) had the fewest nucleotide differences from the wild-type sequence (0 to 2 differences). In comparison, the clones from the AG0.10-CT0.20 and AG0.07-CT0.20 imbalanced Ep-PCRs displayed larger numbers of nucleotide differences (7 to 19 and 7 to 23, respectively). Table 5 shows the genetic diversity captured by PAQ for each of the mutant libraries. Apart from the differences in number of nucleotide differences among the clones, PAQ did not identify any clusters of variants that were similar to each other. We designated the RNA library generated by AG0.20-CT0.20 PCR "cloud-small" (Cloud^{s}), the AG0.10-CT0.20 imbalanced EP-PCR library "cloud-moderate" (Cloud^{m}), and the AG0.07-CT0.20 imbalanced EP-PCR library "cloud-large" (Cloud^{l}). An RNA library prepared using an equimolar (total of 15 μg) mixture of the above three balanced and imbalanced Ep-PCR libraries was designated "cloud-variable" (Cloud^{v}). The wild-type RNA library was designated "cloud-null" (Cloud^{null}) (Table 5).

TABLE 5 PAQ-captured genetic diversity among different mutant Ep-PCR libraries and mutant cloud designations

RNA library	PAQ-captured genetic diversity of variants (no. of mutations) (range [median, mean])	Mutant cloud designation (abbreviation)
Wild type	0 (0, 0)	Cloud-null (Cloud^{null})
Balanced Ep-PCR (0.20 mM [each] dATP, dGTP, dCTP, and dTTP)	0–2 (1, 1.0)	Cloud-small (Cloud^{s})
Imbalanced (0.20 mM [each] dCTP and dTTP and indicated concn [mM] [each] of dATP and dGTP)		
0.10	7–19 (13, 13.3)	Cloud-moderate (Cloud^{m})
0.07	7–23 (17, 16.7)	Cloud-large (Cloud^{l})
Avg range of mutations	3.5–11.0	Cloud-variable (Cloud^{v})

Cloud^{v} promoted transient HEV replication. Variation in the size of a mutant cloud may influence intramutant interactions within the cloud and consequently determine the phenotypic behavior of the virus population in the cloud. Cloud^{v} is a mixture of equal amounts of WT and balanced and imbalanced Ep-PCR library transcripts, which gave us the opportunity to explore distinct regions of the fitness landscape. To determine whether the cloud size can influence the replicative capacity of the virus population, we quantified HEV RNA accumulated in culture supernatants as well as that within the cells until 30 days posttransfection (p.t.) by using real-time PCR.

Significant amounts of RNA accumulated in Cloud^{v}-transfected cell culture supernatants by day 3 p.t., the earliest time point analyzed; the levels showed a slight increase on day 6 p.t. and persisted on day 12 p.t. In contrast, cells transfected with Cloud^{l}, Cloud^{m}, or Cloud^{s} produced at least 2 log lower RNA levels; also, the HEV RNA levels in these cells progressively decreased over time, and no RNA was detectable by day 12 p.t. HEV RNA was not detectable in Cloud^{null} and Sar55 GAD culture supernatants. RNA levels generated by Cloud^{v}-transfected cells were significantly higher than those in Cloud^{null}-, Cloud^{s}-, Cloud^{m}-, or Cloud^{l}-transfected cells on day 6 p.t. ($P < 0.001$; Student *t* test) (Fig. 2A).

For intracellular viral RNA levels, no significant difference was observed between Cloud^{v}-, Cloud^{m}-, Cloud^{s}-, and Cloud^{null}-transfected cells (~5 to 6 log each) at day 3 p.t. There was a transient increase in RNA levels in Cloud^{v}-transfected cells on day 6 p.t. In contrast, in cells transfected with Cloud^{l}, Cloud^{m}, Cloud^{s}, or Cloud^{null}, the amount of RNA declined progressively from day 3 p.t. to day 12 p.t. (Fig. 2B). Intracellular RNA was detectable until day 12 p.t. in Cloud^{v}-transfected cells (data not shown). All HEV-transfected cultures were monitored for extracellular and intracellular viral RNAs until day 30 p.t. by 3-day interval analysis, and no detectable RNAs were observed (data not shown).

Cloud^{v}-transfected cells expressed significant levels of ORF1 protein and dsRNA. Next, we studied the impact of cloud size on the cellular content of the HEV ORF1 protein and double-stranded RNA (dsRNA), a marker of virus replication, by using immunofluorescence. We analyzed the ORF1 protein and dsRNA in S10-3 cells which were either mock transfected or HEV transfected, fixed on day 6 p.t., and then incubated first with ORF1 polyclonal antiserum or monoclonal antibody (MAb) J2 and then with a secondary antibody. The specific fluorescence signal for the ORF1 protein was readily observed in all HEV-transfected cells, but no positive signal was observed in mock-transfected cells (Fig. 3A). Densitometry analysis showed that the ORF1 content in Cloud^{v}-transfected cells was significantly higher ($P < 0.001$; Student *t* test) than that in the cells transfected with Cloud^{l} or Cloud^{null} (Fig. 3C). No significant difference in ORF1 content was observed between cells transfected with Cloud^{l} and those transfected with Cloud^{null}. We could not estimate the infectivity of mutant clouds in S10-3 cells because the cells showed good transfection efficiency; a similar observation was also reported earlier (29).

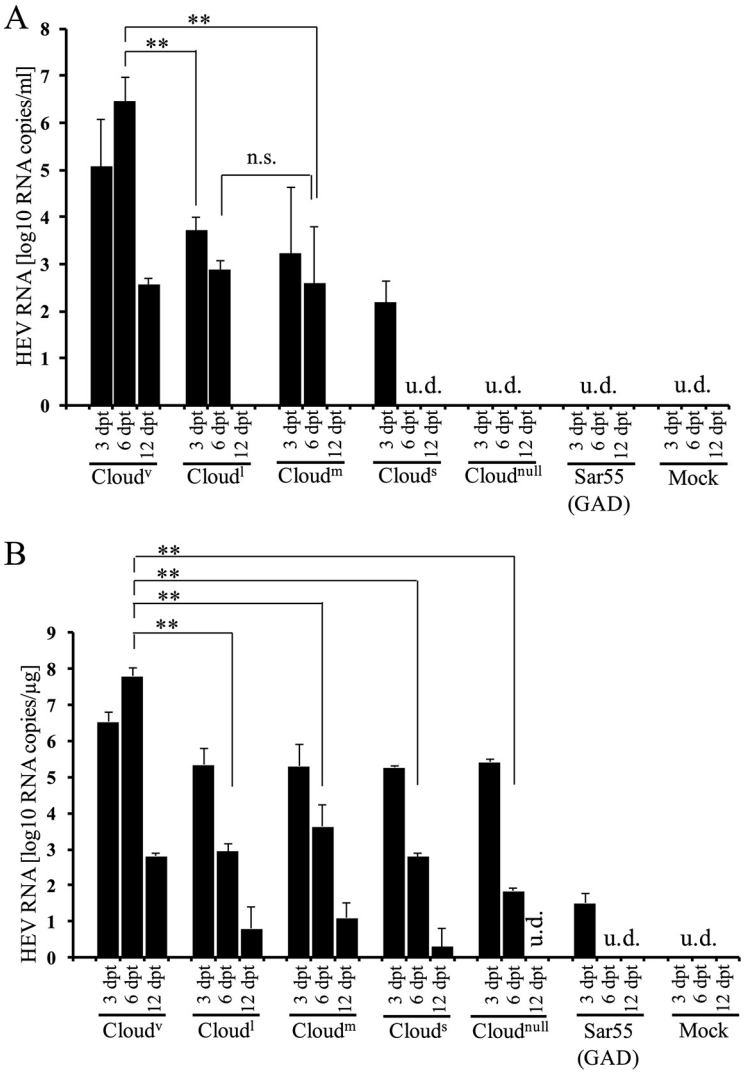

FIG 2 Cloud^{v}-transfected cells have enhanced generation of extracellular and intracellular viral RNAs. Accumulated viral RNA was measured by using quantitative RT-PCR at 3-day intervals for HEV-transfected or mock-transfected cells. HEV RNA levels ($\log_{10}$ RNA copies per milliliter of culture supernatant [A] or per microgram of intracellular RNA [B]) are shown. Results are representative of three independent experiments, each performed in triplicate, and are means and standard deviations (SD). Statistically significant differences are indicated by asterisks (**, $P < 0.001$; Student *t* test). n.s., not significant; u.d., undetectable; dpt, days posttransfection.

HEV replicates in the cytoplasm, similar to most other RNA viruses. Virus dsRNA is a replication intermediate which is formed during the synthesis of nascent ss(+) RNA genomes. It has not been shown that HEV produces dsRNA during the replication process. Using MAb J2, which specifically recognizes dsRNAs, it has been shown that for several RNA viruses, the presence of a dsRNA signal is dependent on replication of the virus genome (30). All HEV-transfected cells showed a strong dsRNA signal (Fig. 3B), but mock-transfected cells did not. On densitometry, the dsRNA signal was stronger in cells transfected with Cloud^{v} ($P < 0.001$; Student *t* test) than in those transfected with Cloud^{l} or Cloud^{null} (Fig. 3D). No significant difference in dsRNA signal was observed between cells transfected with Cloud^{l} and those transfected with Cloud^{null}. The amounts of dsRNA within the cells transfected with Cloud^{v} correlated with the levels of RNA synthesis, as determined by real-time PCR, and with the amount of ORF1 protein accumulated at day 6 p.t. By using MAb J2, an antibody that recognizes dsRNA formed during replication, we were able to provide additional evidence beyond production of

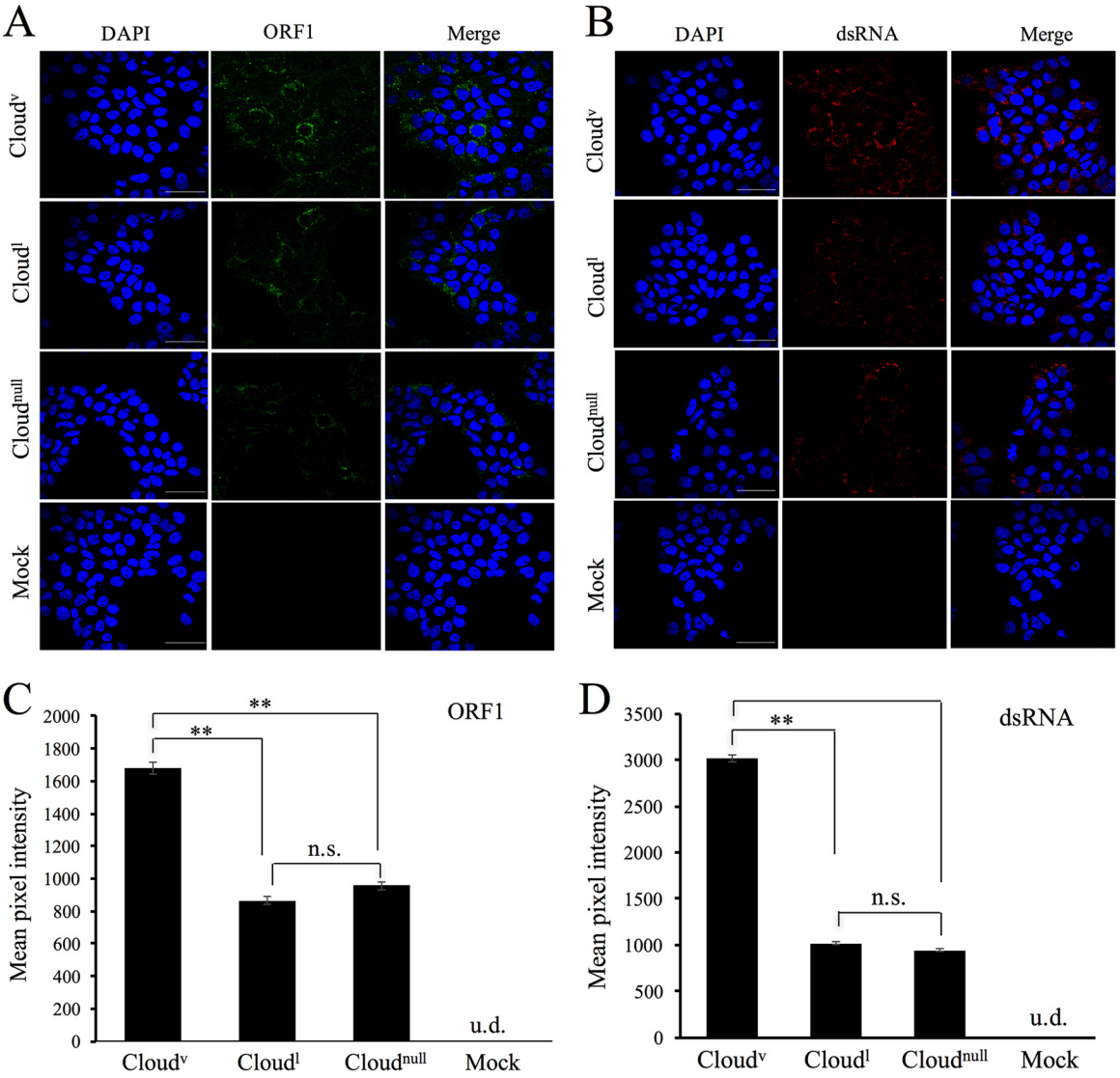

FIG 3 Cloud^{v}-transfected cells have enhanced amounts of ORF1 protein and dsRNA. S10-3 cells at day 6 posttransfection were fixed and stained with mouse polyclonal antibodies to ORF1 (A) or with a rabbit MAb to dsRNA (B), followed by secondary staining with a goat anti-mouse or goat anti-rabbit antibody conjugated to Alexa Fluor 488 or 568 and with DAPI. Bar, 50 μm (applies to all panels). Results are representative of three independent experiments. Densitometry quantitation results for ORF1 (C) and dsRNA (D) are expressed as mean pixel densities. Cloud^{v} significantly increased ORF1 protein and dsRNA levels. Data are presented as means and SD for multiple areas. Statistically significant differences are indicated by asterisks (**, $P < 0.001$; Student t test [95% confidence interval]). n.s., not significant; u.d., undetectable.

ORF1 protein for comparing the effects of various mutant clouds on HEV replication. Based on these data, we can conclude that virus variants in mutant clouds have the capacity to self-replicate.

Cloud^{v}-transfected cells produced infectious virus. To determine whether the S10-3 cells transfected with Cloud^{v} and Cloud^{l} released infectious virus, we infected naive S10-3 cells with supernatants collected at day 6 p.t. Immunofluorescence staining at 5 days postinfection (p.i.) showed ORF1-positive cells in the culture (Fig. 4), which allowed us to determine the number of focus-forming units (FFU) per milliliter of supernatant. We found that Cloud^{v} released 7 FFU/ml on day 6 p.t. However, when the supernatants were diluted 1:10, the infection did not result in discrete foci of ORF1-positive cells (data not shown). Infection with supernatant collected from Cloud^{l}-transfected (day 6) cultures also did not result in discrete foci of ORF1-positive cells (Fig. 4). This finding may have resulted from an inability of the virus released from Cloud^{l}-

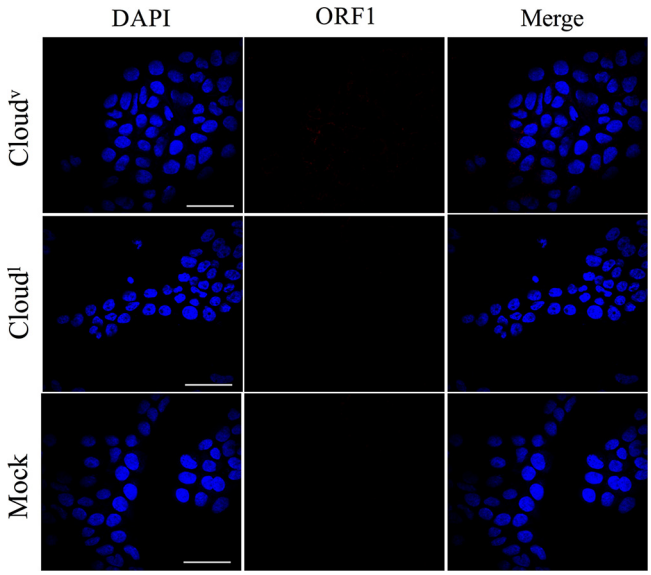

FIG 4 Viruses produced by Cloud^{v}-transfected cells were infectious. Naive S10-3 cells were infected with the viruses produced by Cloud^{v}-, Cloud^{l}-, and mock-transfected cells at day 6 p.t. Cells at day 5 postinfection were fixed and stained with mouse polyclonal antibodies to ORF1, followed by secondary staining with a mouse anti-mouse or goat anti-rabbit antibody conjugated to Alexa Fluor 568 and with DAPI. Bars, 50 μm (applies to all panels). Results are representative of three independent experiments.

transfected culture to enter the S10-3 cells or an inability to establish successful infection in S10-3 cells.

Viruses released from Cloud^{v}-transfected cells were a mixture of unique variants. We asked whether the virus released by cells transfected with Cloud^{v} consisted of a mixture of genetically distinct variants or one or a few predominant variants. To answer this question, we amplified nucleotides 1 to 4936 of the *ORF1* gene in three fragments. Fragment sizes, their corresponding nucleotide and amino acid numbering, and the number of nucleotides sequenced for each fragment are shown in Table 6. All clones of the three fragments had distinct sequences. All putative adaptive mutations (100%) in the clones were substitutions (Fig. S2 to S4). The ratios of transitions to transversions for fragments 1, 2, and 3 were 2.7, 2.4, and 1.0, respectively, and the *dN*/*dS* ratios for fragments 1, 2, and 3 were 1.17, 1.13, and 2.82, respectively (Table 6). The average *dN*/*dS* ratio for all three fragments taken together was 1.62 (Table 6), which is close to the average *dN*/*dS* ratio for the balanced and imbalanced Ep-PCR libraries (Table 2). Because of either low sensitivity of the primers or insufficient levels of virus produced by Cloud^{l}, Cloud^{s}, Cloud^{m}, or Cloud^{null} on day 6 p.t., we were unable to amplify the HEV genome.

Cloud^{v}-transfected cells had higher ISG induction levels. We next asked whether large amounts of dsRNA, a pathogen-associated molecular pattern (PAMP), in Cloud^{v}-transfected cells were associated with a change in the magnitude of induction of ISGs, a widely used surrogate marker for virus infection in experimental settings (31). To investigate this, we compared the expression of ISGs in Cloud^{v}-transfected cells to that in Cloud^{null}-transfected cells, using mock-transfected cells as a control (Fig. 5). Compared to cells transfected with Cloud^{null}, Cloud^{v}-transfected cells showed significant increases in mRNA levels for the ISGs *OAS1* (5-fold increase) (Fig. 5A), *Viperin* (6-fold increase) (Fig. 5B), *RIG-I* (4-fold increase) (Fig. 5C), and *IFITM3* (1.5-fold increase) (Fig. 5D). The Cloud^{null}-transfected cells, in turn, had 4- to 23-fold increases in expression of ISGs compared to the levels in mock-transfected cells (Fig. 5A to D).

DISCUSSION

An understanding of the phenotypic behavior of mutant clouds of RNA viruses in cell culture model systems is essential for understanding the pathogenesis of ss(+) RNA

TABLE 6 Putative adaptive mutations determined by Sanger sequencing

HEV fragment[a]	No. of nucleotides analyzed	No. of mutations (frequency)[b,c]	No. of transitions/no. of transversions (ratio)	No. of nonsynonymous mutations/no. of synonymous mutations (ratio)	Putative adaptive mutations[d]
1 (1.3 kb, nt 32–1360, aa 3–445)	9,971	28 (2.8×10^{-3})	19/7 (2.7)	14/12 (1.17)	S → T (170), V → A (196), V → E (353), R → C (420), R → H (180), Y → F (385), D → G (248), H → R (249), N → D (73), V → D (115), Y → N (221), N → K (129), M → I (178), G → S (425)
2 (1.9 kb, nt 1361–3250, aa 446–1075)	15,128	34 (2.2×10^{-3})	24/10 (2.4)	18/16 (1.13)	F → L (451), C → R (471), Y → C (532), S → G (698), Q → R (773), D → G (810), L → R (948), V → A (976), C → R (929), S → C (898), A → T (602), E → G (724), P → S (1059), F → L (1063), F → I (594), K → R (570), I → K (934), Q → H (1022)
3 (1.68 kb, nt 3251–4936, aa 1076–1637)	13,454	43 (3.2×10^{-3})	21/21 (1.0)	31/11 (2.82)	N → Y (1226), L → M (1245), E → G (1392), I → T (1394), Q → K (1164), M → L (1355), I → M (1496), M → K (1531), G → D (1188), N → Y (1222), I → N (1239), K → R (1387), I → K (1554), L → Q (1306), A → E (1588), F → Y (1605), N → S (1613), R → C (1302), L → H (1366), L → S (1473), L → S (1577), A → P (1606), N → K (1526), T → A (1156), E → G (1192), L → H (1309), P → S (1262), I → F (1338), V → A (1443), P → Q (1505), T → S (1610)
Total			64/38 (1.68)	63/39 (1.62)	

[a]Eight clones were analyzed for each fragment. nt, nucleotide; aa, amino acid.
[b]The number of total mutations was determined relative to the wild-type Sar55 genome sequence (GenBank accession number AF444002) for sequences of the corresponding amplicons.
[c]Mutation frequency is the total number of mutations found divided by the total number of nucleotides sequenced.
[d]Numbers indicate amino acid positions, numbered according to the Sar55 ORF1 protein sequence (GenBank accession number AF444002).

viruses. The ability to induce mutagenesis, albeit in a controlled fashion, in the FL genomes of RNA viruses has long been sought as a model molecular system for investigating mutant clouds in cultured cells. Here we describe the development of a convenient and rapid system for generating HEV RNA mutant clouds which were then delivered into hepatoma cells as a nonselected input pool for phenotypic analysis.

In vitro mutagenesis approaches have been categorized as non-PCR-based and PCR-based approaches. Both these techniques require the *in vitro* generation of a mutagenized gene of interest as a first step. Non-PCR-based *in vitro* mutagenesis approaches demand in-depth mechanics and are difficult to implement as a general strategy (reviewed in reference 32). PCR-based methods include the recently introduced error-prone rolling circle amplification (Ep-RCA) method (33) and the more widely employed Ep-PCR method (20). Ep-RCA products are concatemers, which are not ideal for generating ss(+) RNA virus mutant clouds. *In vitro* mutagenesis of a target gene by use of Ep-PCR is usually followed by digesting the ends of DNA with restriction nucleases, ligating the fragments into the appropriate plasmid vector, and then transforming the products into a suitable host cell to obtain a mutant library. Furthermore, a large pool of transformants is required to increase the diversity of the library, rendering the approach tedious and inefficient.

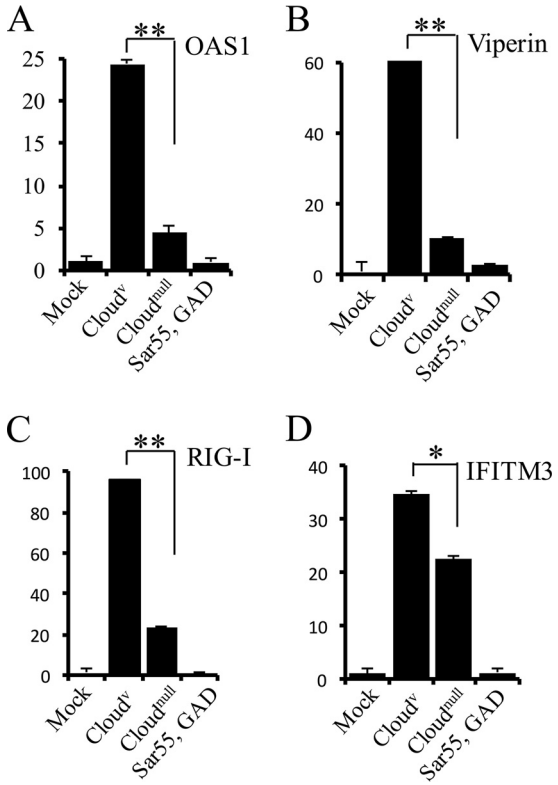

FIG 5 Cloud^{v}-transfected cells were associated with higher ISG induction levels. Induced ISG mRNAs were measured by using quantitative RT-PCR on day 5 p.t. Fold changes for OAS1 (A), viperin (B), RIG-I (C), and IFITM3 (D) mRNAs are shown. Results are representative of three independent experiments, each performed in triplicate, and are means and SD. Statistically significant differences are indicated by asterisks (**, $P < 0.001$; *, $P < 0.05$; Student *t* test [95% confidence interval]). n.s., not significant.

Ep-PCR, our method of choice, uses conditions which, by reducing the fidelity of DNA polymerase, allow incorporation of point mutations during PCR amplification of a full-length genome. In our experiments, the mutation frequency was increased by interfering with the balance of the available dNTPs. Erroneous full-length genome templates are used to synthesize FL RNA genomes with ease because transcription starts by synthesizing the first rGTP (+1 dCTP, located immediately downstream of the T7 RdRp promoter sequence) and "runs off" with rAMP 7202 (+7202 dTMP, derived from the reverse primer). This may ensure that the RNA synthesis reaction produces virus transcripts of defined length. Unlike other bacterium-free approaches, FL-MRC does not require sequential overlap extension PCRs (34) or circular polymerase extension cloning (35) before transfection.

The observed mutation frequency in our experiments was at the upper end of the estimated range of the *de novo* mutation rates for ss(+) RNA viruses, i.e., 10^{-3} to 10^{-5} substitution per nucleotide per replication cycle, consistent with reports of natural evolutionary mutation rates of HEV and hepatitis C virus (HCV) (36). Sequence data demonstrated three important advantages of FL-MRC: (i) the levels of genetic diversity could be controlled; (ii) under conditions of moderate imbalance in the available dNTP pool, a large proportion of induced mutations led to viable proteins; and (iii) a dATP and dGTP imbalance gave a strong bias toward nonsynonymous amino acid changes. These three advantages are strongly desired for the study of the biological changes or evolutionary process of viruses.

None of the putative adaptive amino acid changes observed in this study have been reported previously, and a majority (95%) of them were not found in HEV genotype 1 isolates (data not shown), indicating that the amino acid changes reported here are a potential source of adaptive mutations for HEV. Cooperative interactions among pu-

tative adaptive amino acids may have promoted virus replication. The mean genetic distance observed for virus variants produced by Cloud^{v} transfection was 0.005 (data not shown), which is similar to the intrahost genetic diversity determined by Cerni et al. by use of single-strand conformation polymorphism analysis of the capsid gene (24).

Our analysis did not account for mutational errors resulting from the *in vitro* transcription step. However, the T7 DNA-dependent RNA polymerase has been reported to have an error rate of only about 6×10^{-6} mutation/nucleotide (37), which is 10- to 1,000-fold lower than the estimated rate of mutation in ss(+) RNA and thus should not have influenced our analysis. PEDEL-AA predicts high library diversity at the protein level and suggests that the Ep-PCR libraries should be highly useful at the functional level. The library sizes estimated in our study (5×10^{6}, on average) are comparable to the reported library size (1.4×10^{7}) of the 657-bp *cynT* gene (38). Full-length genome sequence data showed that 2 of 4 variants (50%) had no stops or deletions in their sequences, which is in agreement with the PEDEL-AA prediction that ~50% of variants in the total library do not have indels/stops.

It is reasonable that PAQ did not identify any distinct quasispecies clusters within the clones. Because the experimental conditions that we used produced mutations at random, there was no selective pressure for functionality, and the mutations occurred in many different and random locations rather than at a fixed location. The mutants created by Cloud^{v} transfection represented a genetically diverse swarm of variants that should be advantageous for studying phenotypic properties, and this experimental methodology may help us to examine virus evolution and the emergence of viral diseases.

FL-MRC allows for investigation of the phenotypes of mutant clouds, which has previously not been done for ss(+) RNA viruses. Mutant cloud diversity, replicative capacity, and the host environment in which the virus is multiplying are interdependent variables which may influence the adaptability of virus populations. Cloud^{v}, with multiple coexisting mutants, seems to indicate the existence of selection pressures (host factors) on the mutant population. The higher replication capacity demonstrated by Cloud^{v} may be explained by the hypothesis that the mutants occupy distinct regions in a fitness landscape, which permits them to replicate. Other possible explanations include the following: (i) the level of genetic diversity and composition of the mutant cloud are sufficient to promote cooperative interactions among ORF1 protein variants, but improved catalytic activity in some ORF1 protein variants cannot be ruled out completely; (ii) ORF2 protein variants may fulfill their primary function of encapsidation; and (iii) molecular intermediates of the replication cycle might have enhanced numbers of functional replication complexes, leading to abundant (+)-RNA progeny. dsRNA, a viral PAMP that induces innate immune pathways, has been used effectively as a marker for active virus infection (39). The correlation between expression of ISG mRNAs and the levels of dsRNA and ORF1 protein implies that replication intermediates from the virus infection cycle actively trigger interferon-dependent or -independent innate immune pathways, or both.

Cell culture systems have been proven to be useful for an approximation and basic understanding of the mutant cloud dynamics, heterogeneity, and fitness landscape of vesicular stomatitis virus (40) and foot-and-mouth disease virus (41). Our study involves artificially elevated mutation frequencies, and we acknowledge that not all RNA sequences generated through FL-MRC are directed to become viable mutants, and thus that our findings may not necessarily reflect the true behavior of the circulating mutant cloud in an infected individual.

The most important finding of this study is that mutant clouds of HEV are not equally adaptive; rather, their behavior varies depending on the cloud size and composition. This may be related to the difference in the abilities of different mutant clouds to overcome the threshold of selection pressures imposed by the host-virus infection cycle. Earlier, for experimental poliovirus infection, it was shown that low-fidelity poliovirus (42) or an artificially expanded cloud of poliovirus (9) was able to overcome the gut-blood barrier to reach the brains of infected mice. Given the short duration of

viremia and the absence of a known animal reservoir for genotype 1 HEV, larger mutant clouds might provide a source of variants which have a selective advantage for transmission to new hosts. Thus, further investigations into the dynamic behavior of HEV quasispecies *in vivo* to increase our understanding of this behavior are warranted.

Advancements in molecular biology and the utility of virus reverse genetics allowed us to rapidly produce mutant clouds for a pathogen of public health importance. This method has the potential to simultaneously generate a wide variety of modified viruses while providing a control on the degree of artificial heterogeneity, something which was previously not possible. The molecular system presented here is readily adaptable to other ss(+) RNA viruses and can be used to study the role of genetic diversity in adaptation and evolution of viruses.

MATERIALS AND METHODS

Plasmids and cell lines. Plasmid pSK-E2, encoding a wild-type HEV genotype 1 strain (Sar55) and its replication-deficient GND mutant (Sar55 GND; used as replication-deficient control), and S10-3, a hepatoma cell line, were kind gifts from Suzanne U. Emerson (National Institutes of Health, Bethesda, MD). The S10-3 cells were cultured at 5% CO_2 and 37°C in Dulbecco's modified Eagle medium (DMEM; Life Technologies, Carlsbad, CA) supplemented with 10% heat-inactivated fetal bovine serum (FBS), 100 U/ml penicillin, and 100 μg/ml streptomycin. The cells were split every second to third day at a ratio of 1:2 to 1:3.

Virus full-length amplicon synthesis using Ep-PCRs and sequencing. The full-length (FL) Sar55 genome was genetically diversified by using Ep-PCR to create mutant clouds. In brief, a 7,321-bp fragment consisting of the FL HEV genome was amplified in a 50-μl reaction mixture containing sense (HEV-Neg96 [5′-GGAAACAGCTATGACCATGATTACGCCAAGC-3′; located 96 nucleotides upstream of the T7 promoter]) and antisense (HEV-Pos7208 [5′-TTTTTTTTTTTTTTTCAGGGAGCGCGAAACGCAGAAAAGAG AAATAAG-3′; located at the end of the 3′ UTR of the HEV genome]) primers at 0.2 μM (each), 2 mM $MgCl_2$, 0.05 mM $MnCl_2$, 2 μl KB extender, 1 U Platinum *Taq* DNA polymerase (Invitrogen), and dinucleotides at imbalanced concentrations, i.e., two dNTPs at a final concentration of 0.20 mM (each) and the other two dNTPs at a lower concentration (0.15 mM, 0.10 mM, 0.07 mM, or 0.05 mM [each]), in a Veriti thermocycler (Applied Biosystems). The cycling conditions were as follows: 1 cycle of 2 min at 94°C; 30 cycles of 20 s at 94°C, 25 s at 60°C, and 7 min at 72°C; and 1 cycle of 10 min at 72°C. The Ep-PCR products were separated in a 0.8% agarose gel, and yields were measured against known amounts of a 1-kb DNA ladder. The amplified DNA was purified using a PCR MinElute extraction kit (Qiagen) and subjected to capillary sequencing to assess mutation frequency.

The column-purified Ep-PCR products generated using 0.07 mM (each) dATP and dGTP and 0.20 mM (each) dCTP and dTTP (AG0.07-CT0.20) were ligated into the pCR-XL-TOPO vector of a TOPO XL PCR cloning kit (Invitrogen). In addition, 3.5-kb fragments generated by digestion of amplicons (generated using 0.10 mM [each] dATP and dGTP and 0.2 mM [each] dCTP and dTTP [AG0.10-CT0.20] or AG0.07-CT0.20) with XbaI and XhoI were gel purified, ligated into the pSK-E2 empty vector (the full Sar55 genome was removed by digestion with XbaI and XhoI), and used to transform *Escherichia coli* DH5α competent cells. Individual colonies obtained were analyzed to confirm the presence of the insert, and the insert was subjected to capillary sequencing. Mutation frequencies were then calculated as the number of mutations per kilobase of amplicon compared to the wild-type Sar55 sequence.

Estimation of Ep-PCR efficiency and diversity in mutant libraries. The number of doublings in Ep-PCR (*d*) was estimated as the log(product/template)/log(2), and PCR efficiency (*eff*) was calculated using the formula $2^{(d/n\ \text{cycles})} - 1$. Library size and protein diversity in Ep-PCR libraries were estimated using an online tool named Programme for Estimating Diversity in Error-Prone PCR Libraries—Amino Acid (PEDEL-AA) (http://guinevere.otago.ac.nz/stats.html), which has been described previously (43). The tool used the following parameters: nucleotide sequence of the template, estimated size of the library, mutation frequency matrix, mean number of mutations per sequence in the library, number of cycles in and efficiency of the Ep-PCR, and mean number of insertions and deletions per sequence. The 5′ UTR sequence data were excluded for calculation of diversity.

Assessment of complexity of mutant libraries. The complexity of the mutant libraries was analyzed using mutation frequency (as described by convention [5], it is not a measure of mutation rate, i.e., misincorporation by the RdRp, but an indication of diversity), normalized Shannon entropy, and Hamming distance. Mutation frequency was calculated as the ratio of the total number of mutations identified to the total number of nucleotides sequenced. The normalized Shannon entropy (S_n) value, an indicator of the heterogeneity of a mutant library, was calculated as follows: $S_n = -\{\sum_i [p_i \ln(p_i)]\}/(\ln N)$, where p_i is the frequency of each sequence in the mutant cloud and N is the total number of sequences analyzed. S_n ranges from 0 to 1, indicating no diversity to maximum diversity, respectively (44).

Phylogenetic analysis. Four full-length sequences (generated using AG0.07-CT0.20 Ep-PCR) were aligned to representative sequences of HEV genotypes 1 to 4, obtained from GenBank, by using the ClustalW algorithm in MEGA7. Phylogeny was inferred by using the neighbor-joining method in MEGA7 (45). Tree topology was assessed by using 100 bootstrap replicates. Genetic diversity among sequence variants of a mutant cloud was quantified as the mean genetic distance calculated for all pairs of nucleotide sequences, using MEGA7. Calculations were performed according to the Kimura two-parameter distance method.

PAQ. The program PAQ uses the Hamming distance as a measure of distance between variants and a nonhierarchical clustering method to identify discrete and cohesive partitions of closely related sequences, similar to other clustering algorithms (46). PAQ analysis allowed us to capture the level of genetic diversity in the mutant RNA libraries and to examine the genetic evolution of the viral clouds. The program is currently limited to sequences of up to 1,500 nucleotides or amino acids. Hence, we could not use the entire sequences (~3,260 nucleotides) and had to reduce the length of the sequences to be analyzed. We used Multiple Align Show, a free online sequence alignment tool (http://www.bioinformatics.org/SMS/multi_align.html), to align the sequences and identify the sites where mutations occurred. Since PAQ examines the number of nucleotide differences among sequences, the sites that are conserved among all sequences do not have an impact on PAQ analysis. Therefore, we deleted the conserved stretches from all the sequences to reduce the lengths to fit under the 1,500-nucleotide limit.

Viral transcript synthesis. Column-purified Ep-PCR products were vacuum concentrated, and the DNA concentration was measured using a spectrophotometer (NanoDrop ND-1000). BglII-linearized (Sar55 and Sar55 GND) plasmid DNA was treated with mung bean nuclease (New England BioLabs) at 30°C for 20 min. *In vitro* transcription was performed in a 10-μl reaction volume, using 1 μg of purified Ep-PCR product or linearized plasmid DNA, 5 mM (each) ATP, CTP, and UTP, 0.5 mM GTP, 3 mM cap analog m^7G (Promega), and T7 enzyme mix (RiboMAX Large Scale RNA production system—T7; Promega), followed by incubation at 37°C for 3 h. The products were digested with RNase-free DNase to remove the DNA template; the transcripts were then purified using an RNA cleanup procedure, and their integrity was analyzed in a 0.8% 3-(*N*-morpholino)propanesulfonic acid formaldehyde denaturing agarose gel.

Transfection of S10-3 cells with *in vitro*-transcribed RNA. RNA transcripts (15 μg) from the WT Sar55 strain or an error-prone mutant library or equimolar mixtures of WT and various error-prone mutant libraries were added to 633 μl Opti-MEM (Invitrogen) and 15 μl Lipofectamine 2000 (Invitrogen). The RNA/Lipofectamine complexes were then added to a T25 flask that contained S10-3 cells at around 80% confluence (achieved 20 to 24 h after seeding) and incubated at 37°C and 5% CO_2 for approximately 12 h. At 12 and 24 h posttransfection (p.t.), the cells were washed with phosphate-buffered saline (PBS). The cells were subcultured, and the culture supernatants/cells were collected at 3-day intervals until 30 days p.t.

Quantitation of HEV RNA. Culture supernatants were concentrated 20-fold by use of 100-kDa-cutoff Vivaspin 20 columns (GE Healthcare). RNA was isolated from 140 μl of concentrated supernatant by use of a QIAamp viral RNA minikit (Qiagen). Cellular RNA was extracted using a Qiagen RNA minikit following the manufacturer's instructions. Virus cDNA was synthesized using a Verso cDNA synthesis kit (Invitrogen) and the JV-Rev primer (47) as the reverse primer in a 20-μl reaction mixture. Each real-time PCR mixture contained 5 μl Power SYBR green master mix (Applied Biosystems, CA), 1 μl primer mix (0.2 μM), and 4 μl cDNA template. The reactions were performed and detected on a StepOnePlus real-time PCR system (Applied Biosystems). The quantity of products generated was measured relative to a standard curve obtained with known amounts of HEV cDNA.

High-fidelity amplification of HEV RNA from culture supernatants. To amplify the HEV genome, virus cDNAs were synthesized by using Superscript III reverse transcriptase and a poly(dT) (25-mer) reverse primer. In a nested PCR procedure, HEV genomes were amplified in four DNA fragments. Primer sequences used to amplify the FL genome have been described previously (29). All high-fidelity PCRs were performed with Phusion high-fidelity DNA polymerase (NEB). To the purified PCR products, a dA overhang was added, followed by TA cloning into pGEM-T Easy (Promega) and Sanger sequencing of 8 clones per fragment.

Confocal microscopy of mutant cloud-transfected cells. On day 3 p.t., the cells were trypsinized and 0.4 million cells transferred to glass coverslips. On day 3 postseeding, the cells grown on glass coverslips were fixed with 100% methanol (for ORF1) or freshly prepared 3% paraformaldehyde (for dsRNA), permeabilized with PBS containing 0.2% Triton X-100 for 5 min, and washed three times for 5 min each in PBS. Nonspecific antigens were then blocked for 1 h by using 1% bovine serum albumin and 5% normal goat serum in PBS with 0.3% Triton X-100. The cells were then incubated with primary antibodies for detection of ORF1 (1:1,000 dilution; polyclonal anti-helicase) or dsRNA (1:500 dilution) (English Scientific) at 4°C overnight and washed three times for 5 min each in PBS. This was followed by incubation with secondary antibodies conjugated to Alexa Fluor 488 (for ORF1) and Alexa Fluor 568 (for dsRNA), both at a 1:1,000 dilution at room temperature for 30 min, counterstaining with 1 μg/ml 4′,6-diamidino-2-phenylindole (DAPI) solution for 3 min, and washing three times with PBS containing 0.05% Triton X-100 for 5 min, twice with PBS for 5 min, and once with distilled H_2O for 1 min. The coverslips were mounted on slides, sealed using nail paint, and analyzed by fluorescence microscopy, using a Nikon A1 spinning-disk confocal microscope. Densitometry analysis of immunofluorescence staining was performed on digitized images.

***In vitro* infectivity assays.** To determine the specific infectivity of HEV particles secreted into the extracellular milieu, naive S10-3 cells at 70% confluence were infected for 24 h with concentrated supernatants containing 6×10^5 virus particles as measured by quantitative reverse transcription-PCR (RT-PCR) (harvested from day 6 p.t. cultures). On day 5 postinfection (p.i.), 400,000 cells were seeded onto coverslips. After 24 h, cells grown on coverslips were fixed with 100% methanol, stained for the HEV ORF1 protein, counterstained with DAPI, and imaged as described above. One focus-forming unit (FFU) was defined as one or more infected cells separated from other infected cells by at least 2 uninfected cells. Wherever FFU were observed, a 10-fold dilution of virus was tested for infectivity.

Quantitative RT-PCR. Cellular RNA (1 μg) was reverse transcribed to cDNAs by use of a Verso cDNA synthesis kit (Invitrogen), using a 20-μl reaction mixture which contained a mixture of poly(dT) and

random hexamers at a 1:3 ratio. Real-time PCRs were performed in triplicate for ISGs to obtain a mean threshold cycle (C_T) value for each sample. The expression level of each gene was normalized to that of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) by using the $2^{m\Delta\Delta CT}$ method. Primer sequences are available upon request.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JVI.01932-17.

SUPPLEMENTAL FILE 1, PDF file, 3.6 MB.

ACKNOWLEDGMENTS

We thank Milan Surjit, Translational Health Sciences and Technology Institute, Faridabad, India, for rabbit anti-ORF1 serum.

This study was supported by a grant to N.S.V. from the Science and Engineering Research Board (grant YSS/2014/000107), Department of Science and Technology, under its Young Scientist Scheme. S.A. was supported by a graduate scholarship from Shiv Nadar University.

No author has any financial, professional, or personal potential conflicts that are relevant to the article.

REFERENCES

1. Ahlquist P, Noueiry AO, Lee WM, Kushner DB, Dye BT. 2003. Host factors in positive-strand RNA virus genome replication. J Virol 77:8181–8186. https://doi.org/10.1128/JVI.77.15.8181-8186.2003.
2. Steinhauer DA, Domingo E, Holland JJ. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 122:281–288. https://doi.org/10.1016/0378-1119(92)90216-C.
3. Sanjuan R, Domingo-Calap P. 2016. Mechanisms of viral mutation. Cell Mol Life Sci 73:4433–4448. https://doi.org/10.1007/s00018-016-2299-6.
4. Domingo E, Martinez-Salas E, Sobrino F, de la Torre JC, Portela A, Ortin J, Lopez-Galindez C, Pérez-Brena P, Villanueva N, Najera R, VandePol S, Steinhauer D, DePolo N, Holland J. 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. Gene 40:1–8.
5. Domingo E, Holland JJ. 1997. RNA virus mutations and fitness for survival. Annu Rev Microbiol 51:151–178. https://doi.org/10.1146/annurev.micro.51.1.151.
6. Forns X, Purcell RH, Bukh J. 1999. Quasispecies in viral persistence and pathogenesis of hepatitis C virus. Trends Microbiol 7:402–410. https://doi.org/10.1016/S0966-842X(99)01590-5.
7. Sanchez G, Bosch A, Gomez-Mariano G, Domingo E, Pinto RM. 2003. Evidence for quasispecies distributions in the human hepatitis A virus genome. Virology 315:34–42. https://doi.org/10.1016/S0042-6822(03)00483-5.
8. Steinhauer DA, Holland JJ. 1987. Rapid evolution of RNA viruses. Annu Rev Microbiol 41:409–433. https://doi.org/10.1146/annurev.mi.41.100187.002205.
9. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439:344–348. https://doi.org/10.1038/nature04388.
10. Domingo E, Sheldon J, Perales C. 2012. Viral quasispecies evolution. Microbiol Mol Biol Rev 76:159–216. https://doi.org/10.1128/MMBR.05023-11.
11. Mas A, Lopez-Galindez C, Cacho I, Gomez J, Martinez MA. 2010. Unfinished stories on viral quasispecies and Darwinian views of evolution. J Mol Biol 397:865–877. https://doi.org/10.1016/j.jmb.2010.02.005.
12. Perales C, Lorenzo-Redondo R, López-Galíndez C, Martínez MA, Domingo E. 2010. Mutant spectra in virus behavior. Future Virol 5:679–698. https://doi.org/10.2217/fvl.10.61.
13. Andino R, Domingo E. 2015. Viral quasispecies. Virology 479–480:46–51. https://doi.org/10.1016/j.virol.2015.03.022.
14. Lauring AS, Andino R. 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog 6:e1001005. https://doi.org/10.1371/journal.ppat.1001005.
15. Stobart CC, Moore ML. 2014. RNA virus reverse genetics and vaccine design. Viruses 6:2531–2550. https://doi.org/10.3390/v6072531.
16. Lindenbach BD, Evans MJ, Syder AJ, Wolk B, Tellinghuisen TL, Liu CC, Maruyama T, Hynes RO, Burton DR, McKeating JA, Rice CM. 2005. Complete replication of hepatitis C virus in cell culture. Science 309:623–626. https://doi.org/10.1126/science.1114016.
17. Ramirez S, Mikkelsen LS, Gottwein JM, Bukh J. 2016. Robust HCV genotype 3a infectious cell culture system permits identification of escape variants with resistance to sofosbuvir. Gastroenterology 151:973.e2–985.e2. https://doi.org/10.1053/j.gastro.2016.07.013.
18. Yun SI, Kim SY, Rice CM, Lee YM. 2003. Development and application of a reverse genetics system for Japanese encephalitis virus. J Virol 77:6450–6465. https://doi.org/10.1128/JVI.77.11.6450-6465.2003.
19. Yuan L, Kurek I, English J, Keenan R. 2005. Laboratory-directed protein evolution. Microbiol Mol Biol Rev 69:373–392. https://doi.org/10.1128/MMBR.69.3.373-392.2005.
20. Cadwell RC, Joyce GF. 1992. Randomization of genes by PCR mutagenesis. PCR Methods Appl 2:28–33. https://doi.org/10.1101/gr.2.1.28.
21. Al-Mawsawi LQ, Wu NC, Olson CA, Shi VC, Qi H, Zheng X, Wu TT, Sun R. 2014. High-throughput profiling of point mutations across the HIV-1 genome. Retrovirology 11:124. https://doi.org/10.1186/s12977-014-0124-6.
22. Wedemeyer H, Pischke S, Manns MP. 2012. Pathogenesis and treatment of hepatitis E virus infection. Gastroenterology 142:1388–1397. https://doi.org/10.1053/j.gastro.2012.02.014.
23. Cao D, Meng XJ. 2012. Molecular biology and replication of hepatitis E virus. Emerg Microbes Infect 1:e17. https://doi.org/10.1038/emi.2012.7.
24. Cerni S, Prpic J, Jemersic L, Skoric D. 2015. The application of single strand conformation polymorphism (SSCP) analysis in determining hepatitis E virus intra-host diversity. J Virol Methods 221:46–50. https://doi.org/10.1016/j.jviromet.2015.04.020.
25. Grandadam M, Tebbal S, Caron M, Siriwardana M, Larouze B, Koeck JL, Buisson Y, Enouf V, Nicand E. 2004. Evidence for hepatitis E virus quasispecies. J Gen Virol 85:3189–3194. https://doi.org/10.1099/vir.0.80248-0.
26. Lhomme S, Abravanel F, Dubois M, Sandres-Saune K, Rostaing L, Kamar N, Izopet J. 2012. Hepatitis E virus quasispecies and the outcome of acute hepatitis E in solid-organ transplant patients. J Virol 86:10006–10014. https://doi.org/10.1128/JVI.01003-12.
27. Todt D, Gisa A, Radonic A, Nitsche A, Behrendt P, Suneetha PV, Pischke S, Bremer B, Brown RJ, Manns MP, Cornberg M, Bock CT, Steinmann E, Wedemeyer H. 2016. In vivo evidence for ribavirin-induced mutagenesis of the hepatitis E virus genome. Gut 65:1733–1743. https://doi.org/10.1136/gutjnl-2015-311000.
28. Aggarwal R, Kini D, Sofat S, Naik SR, Krawczynski K. 2000. Duration of viraemia and faecal viral excretion in acute hepatitis E. Lancet 356:1081–1082. https://doi.org/10.1016/S0140-6736(00)02737-9.
29. Chatterjee SN, Devhare PB, Lole KS. 2012. Detection of negative-sense RNA in packaged hepatitis E virions by use of an improved strand-specific reverse transcription-PCR method. J Clin Microbiol 50:1467–1470. https://doi.org/10.1128/JCM.06717-11.

30. Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. 2006. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J Virol 80:5059–5064. https://doi.org/10.1128/JVI.80.10.5059-5064.2006.
31. Marukian S, Andrus L, Sheahan TP, Jones CT, Charles ED, Ploss A, Rice CM, Dustin LB. 2011. Hepatitis C virus induces interferon-λ and interferon-stimulated genes in primary liver cultures. Hepatology 54:1913–1923. https://doi.org/10.1002/hep.24580.
32. Packer MS, Liu DR. 2015. Methods for the directed evolution of proteins. Nat Rev Genet 16:379–394. https://doi.org/10.1038/nrg3927.
33. Fujii R, Kitaoka M, Hayashi K. 2004. One-step random mutagenesis by error-prone rolling circle amplification. Nucleic Acids Res 32:e145. https://doi.org/10.1093/nar/gnh147.
34. Lin TY, Dowd KA, Manhart CJ, Nelson S, Whitehead SS, Pierson TC. 2012. A novel approach for the rapid mutagenesis and directed evolution of the structural genes of West Nile virus. J Virol 86:3501–3512. https://doi.org/10.1128/JVI.06435-11.
35. Edmonds J, Van Grinsven E, Prow N, Bosco-Lauth A, Brault AC, Bowen RA, Hall RA, Khromykh AA. 2013. A novel bacterium free method for generation of flavivirus infectious DNA by circular polymerase extension reaction allows accurate recapitulation of viral heterogeneity. J Virol 87:2367–2372. https://doi.org/10.1128/JVI.03162-12.
36. Ribeiro RM, Li H, Wang S, Stoddard MB, Learn GH, Korber BT, Bhattacharya T, Guedj J, Parrish EH, Hahn BH, Shaw GM, Perelson AS. 2012. Quantifying the diversification of hepatitis C virus (HCV) during primary infection: estimates of the in vivo mutation rate. PLoS Pathog 8:e1002881. https://doi.org/10.1371/journal.ppat.1002881.
37. Sote S, Kleine S, Schlicke M, Brakmann S. 2011. Directed evolution of an error-prone T7 DNA polymerase that attenuates viral replication. Chembiochem 12:1551–1558. https://doi.org/10.1002/cbic.201000799.
38. Hanson-Manful P, Patrick WM. 2013. Construction and analysis of randomized protein-encoding libraries using error-prone PCR. Methods Mol Biol 996:251–267. https://doi.org/10.1007/978-1-62703-354-1_15.
39. Kovalev N, Pogany J, Nagy PD. 2014. Template role of double-stranded RNA in tombusvirus replication. J Virol 88:5638–5651. https://doi.org/10.1128/JVI.03842-13.
40. Novella IS. 2003. Contributions of vesicular stomatitis virus to the understanding of RNA virus evolution. Curr Opin Microbiol 6:399–405. https://doi.org/10.1016/S1369-5274(03)00084-5.
41. Domingo E, Ruiz-Jarabo CM, Arias A, Garcia-Arriaza JF, Escarmis C. 2004. Quasispecies dynamics and evolution of foot-and-mouth disease virus, p 261–304. *In* Sobrino F, Domingo E (ed), Foot-and-mouth disease. Horizon Bioscience, Wymondham, United Kingdom.
42. Pfeiffer JK, Kirkegaard K. 2005. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. PLoS Pathog 1:e11. https://doi.org/10.1371/journal.ppat.0010011.
43. Firth AE, Patrick WM. 2008. GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries. Nucleic Acids Res 36(Web Server issue):W281–W285. https://doi.org/10.1093/nar/gkn226.
44. Shannon C. 1948. A mathematical theory of communication. Bell Syst Tech J 27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
45. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054.
46. Baccam P, Thompson RJ, Fedrigo O, Carpenter S, Cornette JL. 2001. PAQ: partition analysis of quasispecies. Bioinformatics 17:16–22. https://doi.org/10.1093/bioinformatics/17.1.16.
47. Jothikumar N, Cromeans TL, Robertson BH, Meng XJ, Hill VR. 2006. A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. J Virol Methods 131:65–71. https://doi.org/10.1016/j.jviromet.2005.07.004.