The Fecal Viral Flora of California Sea Lions^{∇}[†]

Linlin Li,^{1,2} Tongling Shan,^{1,3} Chunlin Wang,⁴ Colette Côté,⁵ John Kolman,⁵ David Onions,⁵ Frances M. D. Gulland,⁶ and Eric Delwart^{1,2*}

Blood Systems Research Institute, San Francisco, California¹; Department of Laboratory Medicine, University of California, San Francisco, California²; Zoonosis and Comparative Medicine Group, Shanghai Jiao Tong University, Shanghai, China³; Stanford Genome Technology Center, Stanford, California⁴; BioReliance, Rockville, Maryland⁵; and The Marine Mammal Center, Sausalito, California⁶

Received 4 May 2011/Accepted 20 July 2011

California sea lions are one of the major marine mammal species along the Pacific coast of North America. Sea lions are susceptible to a wide variety of viruses, some of which can be transmitted to or from terrestrial mammals. Using an unbiased viral metagenomic approach, we surveyed the fecal virome in California sea lions of different ages and health statuses. Averages of 1.6 and 2.5 distinct mammalian viral species were shed by pups and juvenile sea lions, respectively. Previously undescribed mammalian viruses from four RNA virus families (*Astroviridae, Picornaviridae, Caliciviridae*, and *Reoviridae*) and one DNA virus family (*Parvoviridae*) were characterized. The first complete or partial genomes of sapeloviruses, sapoviruses, noroviruses, and bocavirus in marine mammals are reported. Astroviruses and bocaviruses showed the highest prevalence and abundance in California sea lion feces. The diversity of bacteriophages was higher in unweaned sea lion pups than in juveniles and animals in rehabilitation, where the phage community consisted largely of phages related to the family *Microviridae*. This study increases our understanding of the viral diversity in marine mammals, highlights the high rate of enteric viral infections in these highly social carnivores, and may be used as a baseline viral survey for comparison with samples from California sea lions during unexplained disease outbreaks.

California sea lions (*Zalophus californianus*) have a population of approximately 240,000 and along with seals and walruses are members of the subgroup Pinnipedia in the suborder Caniformia in the order Carnivora. They inhabit mainland shorelines and coastal islands along the west coast of North America and migrate along the coast during the nonbreeding season. California sea lions are strict carnivores, eating a variety of marine prey, including more than 50 species of fishes and cephalopods. Sea lion pups start eating fish at about 5 months of age, in addition to their mother's milk; are weaned at 10 to 12 months old; and can live up to 15 to 25 years. California sea lions are gregarious animals, forming large rookeries at breeding sites, and aggregate at high densities on haul-out sites (7).

California sea lions share beaches and coastal waters with humans, often resting on human-made structures such as docks and boats, and are affected by pathogens and chemicals that enter coastal waters through runoff and sewage outfalls (5). Features of California sea lions, including their large population, wide geographic distribution and migration, gregarious nature, long life span, and shared environment with humans, may favor the transmission of viruses among themselves and to and/or from humans and other mammals.

A commonly reported sea lion virus is San Miguel sea lion virus (SMSV), a calicivirus in the genus *Vesivirus*. SMSV was first isolated from California sea lions from San Miguel Island

in 1972 (53). SMSV causes vesicular lesions of the skin and mucosa, abortion, pneumonia, and encephalitis in sea lions and is transmissible to swine, generating a disease identical to that caused by vesicular exanthema of swine virus (VESV), a very closely related calicivirus (43, 65). The epidemics of VESV in North America from the 1930s to 1950s were shown by classical virological investigations to be serotypes of marine origin (56, 57, 65). SMSV was also found in vesicular lesions in humans, and antibodies were detected in blood donors (54, 55).

Canine distemper virus (CDV), a paramyxovirus of the genus *Morbillivirus*, was first described in 1905 (48). CDV most commonly affects dogs and causes gastrointestinal and respiratory symptoms as well as neurological symptoms. It also infects other domestic and wild carnivores, including ferrets, mink, foxes, and raccoons (38). CDV infection is not confined to terrestrial host species and has caused significant problems in marine mammals in the past 2 decades. It was identified as the cause of death of several thousand Baikal seals (*Phoca sibirica*) in 1988 (22) and 10,000 Caspian seals (*Phoca caspica*) in 2000 (30). CDV was also detected in the brain tissue of one captive California sea lion that died of unknown causes in 1995 in Europe (4).

Besides SMSV and CDV, viruses detected in California sea lions also include astrovirus (AstV) (50), polyomavirus (15), anellovirus (45), gammaherpesvirus (10, 32, 36), parapoxvirus (46, 47), retrovirus (31), and adenovirus (18). Otarine herpesvirus 1 is considered a possible contributing factor to the unusually high occurrence of tumors in California sea lions (10, 32, 44). A parapoxvirus was isolated from cutaneous nodular lesions, and the prevalence of antiparapoxviral antibodies in 761 free-ranging California sea lions was 91% (46, 47).

^{*} Corresponding author. Mailing address: Blood Systems Research Institute, 270 Masonic Ave., San Francisco, CA 94118. Phone: (415) 923-5763. Fax: (415) 567-5899. E-mail: delwarte@medicine.ucsf.edu.

[†] Supplemental material for this article may be found at http://jvi .asm.org/.

^v Published ahead of print on 27 July 2011.

In this study, we used next-generation sequencing to get a comprehensive view of the fecal viral populations from wild and temporarily captive California sea lions. We report previously uncharacterized California sea lion viruses, including astroviruses, picornaviruses, bocaviruses, sapoviruses, and other viruses. These results provide a baseline for the current enteric viral burden in this marine mammal species that can be compared to later virome surveys to detect alterations associated with changes in their health or population size.

MATERIALS AND METHODS

Animal specimen collection. Fecal specimens from three groups of California sea lions were collected by The Marine Mammal Center (TMMC) during July to October 2010. A total of 47 fecal specimens were analyzed. One group consisted of feces from 14 pups (3 months old) on San Miguel Island, CA. A second group consisted of feces from 19 juveniles (2 to 3 years old) also on San Miguel Island, CA. A third group consisted of samples from 14 California sea lions (>1 year old) being rehabilitated at TMMC for reasons including malnutrition, domoic acid toxicosis, leptospirosis, cancer, pneumonia, entanglement, and trauma (23). Fecal specimens were stored in small plastic bags and frozen at -80° C. Sample collection was performed under Marine Mammal Protection Act (MMPA) permit no. 932-1905-00/MA-009526 while animals were handled for veterinary examination.

Sample preparation and viral nucleic acid extraction. Fecal samples were processed as previously described (62). Briefly, fecal samples were resuspended by vigorous vortexing in Hanks' buffered saline solution (Gibco BRL) at a concentration of ~0.5 g/ml. The stool suspension was then centrifuged at 10,000 × g for 3 min, and the supernatant was filtered through a 0.45-µm filter (Millipore) to remove bacteria or cellular debris. The viral-particle containing filtrates were digested with a mixture of DNases and RNases to remove unprotected (not in viral capsids) nucleic acids (1). Viral RNA and DNA were then extracted by using the QIAamp viral RNA Minikit (Qiagen). Extracted viral nucleic acids were protected from degradation by the addition of 40 U of RNase inhibitor (Fermentas) and stored at -80° C for future use.

Library construction and pyrosequencing. Combined viral RNA and DNA libraries were constructed by random PCR amplification using reverse transcription (RT) and PCR primers with degenerate 3' ends as previously described (62). Random PCR products were pooled and separated on a 2% agarose gel, and DNA fragments from 500 bp to 1,000 bp were excised and extracted. The resulting purified product was prepared for sequencing by use of a GS FLX Titanium general library preparation kit (454 Life Science, Roche), and the library of single-stranded DNA fragments was sequenced on a single pyrose-quencing gasket by use of a Genome Sequencer FLX instrument (454 Life Science, Roche), generating approximately 616,000 high-quality nucleotide sequence reads with an average length of 260 bases.

Bioinformatics. The pyrosequencing reads were sorted into their fecal samples of origin according to their unique sequence tag (20 fixed bases of the random PCR primer). Primer sequences plus the adjacent 8 nucleotides were then trimmed from each read. Trimmed reads from each sample were assembled *de novo* by using the Mira assembly program (13), with a criterion of 95% identity or greater over \geq 35 bp. The sequences greater than 100 bp were compared to the GenBank nonredundant nucleotide and protein databases using BLASTn and BLASTx, respectively. Sequences were classified into eukaryotic viruses, phages, bacteria, and eukaryotes based on the taxonomic origin of the best-hit sequence. An E value of 0.001 was used as the cutoff value of significant hits.

Phylogenetic analysis. Reference viral sequences from different viral families were obtained from GenBank. Amino acid sequence alignments were generated by using ClustalW, implemented in MEGA 4 with the default settings (33). Aligned sequences were trimmed to match the genomic regions of the viral sequences obtained in this study, and phylogenetic trees were generated by MEGA4 using the neighbor-joining method with amino acid *p*-distances and 1,000 bootstrap replicates. The GenBank accession numbers of the viral sequences used in the phylogenetic analyses are shown in the trees.

Nucleotide sequence accession number. High-quality sequences and contigs of veterinary sample metagenomes have been deposited in the short-read archive of GenBank under accession no. SRA044033.

RESULTS

Virome overview. Viral nucleic acids enriched from 47 fecal samples from California sea lions were randomly amplified and pyrosequenced, generating >600,000 sequence reads. For each sample, sequence contigs were assembled and, together with singlets longer than 100 bases, were taxonomically classified based on the best BLAST scores (E < 0.001). Approximately 25% of the total reads showed detectable similarity to eukaryotic viral sequences, and 21% matched bacteriophage sequences.

The majority of the eukaryotic viruses detected belonged to DNA and RNA viral species not previously reported for marine mammals (Table 1). The most abundant eukaryotic viruses were bocaviruses, from the family *Parvoviridae* (39% of the total eukaryotic virus reads), followed by astroviruses, from the *Astroviridae* (30%); densoviruses, from the *Parvoviridae* (10%); dependoviruses, from the *Parvoviridae* (9%); caliciviruses, from the *Caliciviridae* (7%); picornaviruses, from the *Picornaviridae* (2%); and rotaviruses, from the *Reoviridae* (2%). The most prevalent viruses were astrovirus, bocavirus, and rotavirus, detected in 51%, 38%, and 28%, respectively, of the animals tested. Averages of 1.6, 2.5, and 2.1 distinct mammalian viruses were identified in the feces of individual pups, juveniles, and animals in rehabilitation, respectively.

Comparison of viruses in different sea lion groups. The fecal viral communities of 14 unweaned pups (3 months old), 19 juveniles (2 to 3 years old), and 14 Californian sea lions in rehabilitation for various symptoms (>1 year old) were then compared (Fig. 1). The percentages of eukaryotic virus reads compared to the total number of sequences read were 12%, 24%, and 35% for pups, juveniles, and animals in rehabilitation, respectively. The percentage of eukaryotic virus reads compared to total virus reads (i.e., including prokaryotic viruses) was lowest for the pups (22%), intermediate for wild juveniles (47%), and highest for the animals in rehabilitation (88%) (Fig. 1A). The most abundant bacteriophage-like sequences found were those with similarities to the major double-stranded DNA podoviruses, siphoviruses, and myoviruses and the single-stranded DNA microviruses. The phage community in the pups had the most complex composition, with siphoviruses, podoviruses, myoviruses, and microviruses at 42%, 20%, 5%, and 26% of the total phage reads, respectively, while the phages in juveniles and animals in rehabilitation were less diverse, with 89% and 73% of their phage reads being related to microviruses (Fig. 1B). Using $\geq 80\%$ of the total viral reads as a criterion, phages dominated in 11/14 (79%) pups, 12/19 (63%) juveniles, and 5/14 (36%) animals in rehabilitation, while eukaryotic viruses dominated in 2/14 (14%) pups, 5/19 (26%) juveniles, and 6/14 (49%) animals in rehabilitation.

For all three animal groups, astroviruses and bocaviruses constituted the majority of the eukaryotic virus community (Fig. 1C). Among eukaryotic virus reads, the percentages for astroviruses and bocaviruses combined were 95%, 68%, and 64% for pups, juveniles, and animals in rehabilitation, respectively. Compared with the other two groups, the juvenile group had higher percentages of dependovirus (23% of the total eukaryotic virus reads) and picornavirus (6%) reads, while the group in rehabilitation had higher percentages of calicivirus (14%) and densovirus (21%) reads (Fig. 1C).

TABLE 1. Summary of mammalian viruses found in California sea lion feces

1203 Pup 15,293 Bocavirus (3) 1207 Pup 5,258 Astrovirus (879), rotavirus (1)	
1207 Pup 5.258 Astrovirus (879), rotavirus (1)	
1209 Pup 3,118 Bocavirus (1)	
1211 Pup 10.315 Bocavirus (505), anellovirus (4)	
1214 Pup 5.751 None	
1215 Pup 7.981 Astrovirus (76), bocavirus (17), sapelovirus (45), picobirna	virus (12)
1218 Pup 10.829 Bocavirus (6.743), dependovirus (1), sapovirus (16)	
1219 Pup 13,525 None	
1222 Pup 6.312 Astrovirus (47), cardiovirus (17)	
1223 Pup 15,554 None	
1230 Pup 8,780 Astrovirus (1), dependovirus (8)	
1234 Pup 9,895 Astrovirus (6,091), bocavirus (35)	
1244 Pup 7.353 Rotavirus (723), anellovirus (1)	
1249 Pup 2.150 Bocavirus (1)	
1125 Juvenile 19.677 Astrovirus (341)	
1136 Juvenile 17,100 Astrovirus (2,500), bocavirus (5,917), dependovirus (5,709)	, sapovirus (79)
1137 Juvenile 10.551 Norovirus (6), rotavirus (916), anellovirus (7)	, (·)
1140 Juvenile 10.463 Astrovirus (313), rotavirus (257), anellovirus (2)	
1141 Juvenile 8430 Astrovirus (133), bocavirus (1), rotavirus (7), anellovirus (7)	3)
1148 Juvenile 9.247 Astrovirus (7.701), rotavirus (4), sanelovirus (2)	-)
1153 Juvenile 14,878 Astrovirus (4,332), bocavirus (6,589), parvovirus (2), rotavi enterovirus/sapelovirus (1,716)	irus (68), anellovirus (2),
1157 Juvenile 8.034 Bocavirus (1), rotavirus (7)	
1162 Juvenile 15,735 Bocavirus (12), enterovirus/sapelovirus (1,318)	
1166 Juvenile 5,151 Sapovirus (1), anellovirus (2), sapelovirus (4)	
1169 Juvenile 9,403 Astrovirus (1,987), dependovirus (1)	
1170 Juvenile 4,376 Astrovirus (212), norvirus (49)	
1174 Juvenile 9,174 Astrovirus (32), parvovirus (2), rotavirus (37), sapelovirus	(1)
1181 Juvenile 5,111 Astrovirus (109), norovirus (1)	
1182 Juvenile 16.827 None	
1185 Juvenile 6,319 Astrovirus (115), bocavirus (1), norovirus (43), rotavirus (5	5)
1187 Juvenile 20,882 Bocavirus (7,779), dependovirus (7,292)	,
1194 Juvenile 19,289 Astrovirus (3)	
1199 Juvenile 17,383 None	
9715 Rehab. juvenile 17.882 Astrovirus (6.450), bocavirus (1), densovirus (612), picobir	navirus (1)
9775 Rehab. yearling 16,798 Bocavirus (1), sapovirus (9,672), cardiovirus (28)	
9795 Rehab. juvenile 9,993 Astrovirus (1,564), enterovirus (5)	
9801 Rehab. adult 17,382 Astrovirus (103)	
9805 Rehab. subadult 13,132 Astrovirus (1), bocavirus (7,889), picornavirus (79), picobir	mavirus (1), hepevirus (1)
9806 Rehab. adult 13,747 Calicivirus (199), picornavirus (8)	
9807 Rehab. adult 7,519 None	
9810 Rehab, yearling 7,133 Rotavirus (3)	
9813 Rehab. adult 11,386 Astrovirus (65), vesivirus (3)	
9814 Rehab. adult 14,920 None	
9816 Rehab, juvenile 14,631 Astrovirus (1).	
9822 Rehab. juvenile 29,857 Astrovirus (8,615), bocavirus (19,140), dependovirus (2), sa picobirnavirus (61)	apovirus (1), rotavirus (4),
9828 Rehab. yearling 18,610 Parvovirus (253), papillomavirus (1)	
9830 Rehab. juvenile 3,497 Rotavirus (76), picobirnavirus (1), asfarvirus (5)	

Astrovirus. Astroviruses are positive single-stranded RNA viruses with a genome of 6.4 to 7.3 kb and have been identified in a wide variety of terrestrial mammals and birds (42). Human astrovirus is a significant cause of acute pediatric gastroenteritis (24). Recently, astroviruses have also been detected in marine mammals, including California sea lion, Steller sea lion (50), bottlenose dolphin, killer whale, and minke whale (64). In this study, astroviruses were present in 24 of the 47 California sea lion samples and had high titers in 9 samples (>500 sequence reads). Sequence assembly within each samples generated 9 near-complete genome sequences, covering 80% to 99% of genome length (GenBank accession no. JN420351 to JN420358). As the astroviruses in samples 9715 and 9795

shared 99% nucleotide similarity with each other, a total of 8 astrovirus species were identified and temporarily named California sea lion astrovirus 4 (Csl AstV4) to Csl AstV11.

As typical mamastroviruses, Csl AstV4 to Csl AstV11 had three putative open reading frames (ORFs) encoding nonstructural proteins with ORF1a and ORF1b and a structural protein with ORF2 (Fig. 2A and Table S1 in the supplemental material). The conserved protease motifs, RNA-dependent RNA polymerase (RdRp) motifs, and a ribosomal frameshift signal (AAAAAAC) in the ORF1a/1b overlap region were found in all Csl AstVs.

To determine the divergence in sequence among Csl AstV species and those of other AstVs, sequence alignments of



FIG. 1. Virome comparisons for three California sea lion groups based on BLASTx comparison to the GenBank nonredundant database (E value of <0.001). (A) Percentage of virus-like sequence reads with similarity to bacteriophages and eukaryotic viruses. (B) Percentage of phage-related sequences in different viral families. (C) Percentage of eukaryotic virus-related sequences in different viral groups.

ORF2 (encoding capsid) and ORF1b (encoding RdRp) were performed, and neighbor-joining trees were generated. The tree for the capsid protein confirmed that Csl AstV4 to Csl AstV11 were novel AstV species, having less than 60% amino acid similarity with the recently characterized Csl AstV1 to Csl AstV3 (50) and other AstVs, showing a high level of diversity among AstVs from a single host species (Fig. 2B). The tree for the RdRp region revealed that California sea lion astroviruses



FIG. 2. (A) Genome organization of California sea lion astroviruses (Csl AstVs). (B) Phylogenetic analysis of California sea lion astroviruses. Trees are based on complete capsid (ORF2) proteins. The novel Csl AstV4 and Cs2 AstV11 are marked by black circles, and the previously reported Csl AstV1 and Csl AstV2 are marked by gray diamonds. nt, nucleotides.

formed three genetic clusters (see Fig. S1 in the supplemental material). The use of available RdRp sequences revealed that Csl AstVs 3, 5, and 11 were the mamastroviruses most closely related to the clade comprised of human AstV1 to AstV8, sharing as high as 92% amino acid similarity in the RdRp region.

Picornavirus. Picornaviruses are small, nonenveloped, positive-sense, single-stranded RNA viruses with a genome size of 7.1 to 8.9 kb, encoding a single polyprotein (58). Here, we found picornavirus sequences in 11 California sea lion samples that were abundant (>1,000 reads) in 2 juvenile samples (samples 1153 and 1162) from San Miguel Island. Sample 1162 contained two distinct picornaviruses, one having 99% nucleotide similarity to the strain in sample 1153. Assembly of the viral reads in samples 1153 and 1162 generated long contigs of 2 distinct picornaviruses, each covering more than 70% of the genome (GenBank accession no. JN420367 and JN420368). In sample 1162, a large 6.5-kb sequence spanned a partial 5' untranslated region (5'UTR); the complete leader protein L, P1, and P2 regions; and a partial P3 region, while in sample 1153, two fragments of 2 kb and 3.9 kb yielded a partial P1 region, the complete P2 region, and a partial P3 region (Fig. 3A).

According to the International Committee on Taxonomy of Viruses (ICTV) (http://www.picornastudygroup.com/definitions/genus_definition.htm), the members of a picornavirus genus should share >40%, >40%, and >50% amino acid similarity in their P1, P2, and P3 regions, respectively. As the P1, P2, and partial P3 regions of the picornavirus in sample 1162 shared 46%, 39%, and 52% amino acid similarities with its closest relative, simian sapelovirus 2, the virus was considered a novel species in the genus *Sapelovirus*. The P1, P2, and P3 regions of the sapelovirus in sample 1153 shared 60%, 70%, and 75% amino acid similarity with its closest relative, the sapelovirus in



FIG. 3. (A) Genome organization of California sea lion sapeloviruses (Csl SaVs). (B) Phylogenetic analysis of the partial P1 region of sapeloviruses, including representative enteroviruses.

sample 1162. Therefore, we temporarily named these two novel picornaviruses California sea lion sapelovirus 1 (Csl SapV1) and Csl SapV2. Known hosts for sapeloviruses therefore include pigs, monkeys, birds, and now Californian sea lions.

The genome organization of Csl SapV1 is typical of a picornavirus, with a single large ORF encoding a near-complete polyprotein of 2,054 amino acids (aa). The polyprotein comprised a putative leader protein; the capsid proteins VP4, VP2, VP3, and VP1; and nonstructural proteins 2A to 2C and 3A to 3D (partially sequenced). The L protein was 94 aa and did not have significant similarity to any other protein. Phylogenetic analyses were performed on the partial P1 region and confirmed that Csl SapVs fell into the sapelovirus genus and were located close to the basal nodes of this clade, with a preferred association with mammalian sapeloviruses (Fig. 3B). Phylogenetic analyses based on the partial 3D region yielded a similar topology (data not shown).

Calicivirus. Caliciviruses are single-stranded, positive-sense, nonenveloped RNA viruses with a genome size of 7.3 to 8.3 kb. The family *Caliciviridae* includes multiple genera, including *Sapovirus, Vesivirus, Norovirus,* and *Lagovirus* (20). Among caliciviruses, only vesiviruses have been reported in marine mammals (41, 53). We detected calicivirus sequences in 10 California sea lion samples. The vesivirus identified in one TMMC rehabilitation sample was San Miguel sea lion virus (SMSV), with approximately 90% nucleotide similarity to a previously sequenced isolate. The other caliciviruses identified here were novel sapoviruses (4 samples) and noroviruses (4 samples).

Sapoviruses (SaVs), previously known as Sapporo-like viruses, are important enteric pathogens that cause diarrhea in humans, pigs, dogs, and mink (14, 39, 61). Sapovirus sequences were found at a high abundance of 9,672 reads in a California sea lion with severe osteomyelitis and nephrolithiasis in rehabilitation at TMMC (sample 9775). One near-complete (7.4-kb) genome from sample 9775 and one partial genome (2.2 kb) from sample 1136 could be assembled (GenBank accession no. JN420369 to JN420370). These sapoviruses were temporarily named California sea lion sapovirus 1 (Csl SaV1) and Csl SaV2.

Csl SaV1 has the typical SaV genome organization, with two



FIG. 4. (A) Genome organization of California sea lion sapovirus 1 (Csl SaV1). (B) Phylogenetic analysis of the VP1 regions of Csl SaV1 and sapoviruses from different genogroups. GenBank accession numbers are shown at the end of the branches. GV, genogroup V; RHDV, rabbit hemorrhagic disease virus.

major ORFs (Fig. 4A). The near-complete ORF1 encodes a polyprotein of 2,259 aa that can be theoretically cleaved into the nonstructural protein NTPase, a 3C-like protease, an RdRp, and a major capsid protein (VP1; 563 aa). ORF2 encoded a minor structural protein (VP2; 167 aa). Csl SaV1 shared as high as 57% amino acid similarity in the polyprotein region with human SaVs, while the VP2 protein showed an amino acid similarity of 65% with human and swine SaVs. The complete VP1 region showed the highest similarity (65% amino acid similarity) with human SaVs. The phylogenetic analysis of the complete VP1 protein of representative SaVs in the GenBank database confirmed that Csl SaV1 belonged to that genus and was most closely related to SaV genogroup V (Fig. 4B). The partial VP1 region of 204 aa from Csl SaV2 shared the highest similarity (72% amino acid similarity) and phylogenetically grouped with human SaVs in genogroup II (see Fig. S2 in the supplemental material).

Low titers of noroviruses (<50 sequence reads) were detected in 4 fecal samples from California sea lions. Norovirus is one of the leading causes of human viral gastroenteritis and a major contributor to cases of food-borne illness worldwide. The contamination of water ecosystems (25, 51, 63) and seafood (17, 60) by norovirus has been widely reported. The norovirus genome has three overlapping ORFs. ORF1 encodes 6 nonstructural proteins, including the viral polymerase RdRp, while ORF2 encodes the capsid protein VP1, and ORF3 encodes a minor structural protein, VP2 (20). The norovirus sequences detected here were small fragments of different genome regions. For sample 1170, phylogenetic analyses based on the amino acid sequences of a 399-bp RdRp region (GenBank accession no. JN420373) (see Fig. S3 in the supplemental material) and a 319-bp VP1 region (GenBank accession no. JN420374) (Fig. 5) showed that the newly discovered

Csl Norovirus (NV) 1170



FIG. 5. Phylogenetic analysis of California sea lion norovirus (Csl NV1170) and representatives from different genogroups based on the amino acid sequence of the partial VP1 region.

sea lion norovirus (Csl NV1170) was most related to genogroup II noroviruses, sharing >70% amino acid similarity.

Rotavirus. Rotaviruses are nonenveloped, double-stranded RNA viruses consisting of 11 genomic segments (0.6 to 3.3 kb) with a total genome size of approximately 18 kb. The genus *Rotavirus* belongs to the family *Reoviridae* and contains 7 serologically distinct species (A to G). Rotaviruses have a triple-layered capsid structure made of 6 proteins (VP1 to VP4, VP6, and VP7). Rotaviruses are the most common cause of acute gastroenteritis in infants and young children and are common in a wide variety of terrestrial mammals and birds (49). A recent study reported that rotavirus antibodies were detected in approximately 20% of the serum samples from Galapagos sea lion pups (n = 125) and Galapagos fur seal pups (n = 22). Rotavirus RNA was detected in 1 out of 18 fecal sample from Galapagos sea lion pups (16).

In this study, rotavirus sequences were identified in 2 out of 14 samples from 3-month-old pups and 11 out of 27 samples from yearling/juvenile sea lions. Sequence assemblies produced fragments of different genomic regions, most of them too short to be phylogenetically informative. The rotavirus sequences from two samples (samples 1137 and 1244, each yielding >500 reads) shared 99% nucleotide similarity with each other. This rotavirus was temporally labeled California sea lion rotavirus 1 (Csl RV1). Phylogenetic analyses were performed on the amino acid sequences of a partial VP4 sequence ($\sim 900 \text{ bp}$) and a partial VP2 sequence (~500 bp) from sample 1137/1244 (GenBank accession no. JN420375 to JN420378). The outer capsid protein, VP4, and inner capsid protein, VP2, have both been used to define the rotavirus group/species genetically (16). Csl RV1 VP4 was closest to adult diarrheal rotavirus J19/B219, with a preferred association with group B lineages (Fig. 6). A similar topology was seen with the partial VP2 region (see Fig. S4 in the supplemental material).

Bocavirus. Members of the genus *Bocavirus* from the family *Parvoviridae* are small, nonenveloped, autonomously replicating, single-stranded DNA viruses with a genome length of about 5.4 kb (59). Bocaviruses were initially discovered in the 1960s with two species, bovine parvovirus and canine minute virus (37). Human bocaviruses are newly recognized human parvoviruses first reported in 2005 (2) and have been associated with respiratory tract disease and, possibly, gastroenteritis. Related species of human bocaviruses in feces have since



FIG. 6. Phylogenetic analysis of California sea lion rotavirus 1 (Csl RV1) and representative rotaviruses from other species based on the amino acid sequence of the partial VP4 region.

been reported and associated with gastroenteritis (3, 26, 29). Recently, bocavirus species were also identified in swine (6, 12), gorilla (27), and chimpanzee (52). Here, bocaviruses were identified in 18 of the 47 California sea lion samples and were abundant in 7 samples (>500 reads). The assembly of each sample generated 6 near-complete genome sequences (5.1 to 5.4 kb) and 1 partial genome (2 kb) (GenBank accession no. JN420360 to JN420366). As the bocaviruses in four samples (samples 1136, 1153, 1187, and 1218) were nearly identical (>99% nucleotide similarity), a total of 4 bocavirus species were identified and temporarily named California sea lion bocavirus 1 (Csl BoV1) to Csl BoV4.

The genome organization of the Csl BoVs was similar to that of the other known bocaviruses (Fig. 7A). It is predicted to contain three major ORFs, encoding the nonstructural proteins NS1 and



FIG. 7. (A) Genome organization of California sea lion bocaviruses (Csl BoVs). (B) Phylogenetic analysis of Csl BoVs and representative bocaviruses based on the complete VP1 protein.

NP1 and the structural protein VP1. The NS1 protein was 793 aa for Csl BoV1, Csl BoV2, and Csl BoV4 and 802 aa for Csl BoV3. Conserved motifs associated with rolling-circle replication, helicase, and ATPase were present in NS1. The NP1 protein encoded by the middle ORF, a unique feature of bocaviruses, was 193 aa in all 4 Csl BoV3. The VP1 protein was 719 aa for Csl BoV1, Csl BoV3, and Csl BoV4 and 718 aa for Csl BoV2. Csl BoVs were most closely related to canine minute virus, showing 54%, 60 to 64%, and 64 to 67% amino acid similarities in the NS1, NP, and VP1 regions, respectively. A phylogenetic analysis of the entire VP1 protein was performed to determine the relationship between Csl BoV1 to Csl BoV4 and other bocaviruses. All Csl BoVs clustered together and were closest to canine minute virus (Fig. 7B).

Dependovirus. The genus Dependovirus belongs to the family Parvoviridae and contains small, nonenveloped, singlestranded DNA viruses with a genome length of about 4.7 kb (59). Dependoviruses are mostly replication-defective adenoassociated viruses (AAVs), but some autonomous avian parvoviruses have also been classified into this genus (66). Dependoviruses were found in human and several other mammalian species, avian species, and amphibian species and are considered commensal viruses (21). Here, AAVs were identified in 6 California sea lion samples and were very abundant (>5,000 reads) in 2 juvenile samples from San Miguel Island. The assembly of each sample generated 2 near-complete genome sequences (~4.4 kb) (GenBank accession no. JN420371 and JN420372), sharing 99% nucleotide similarity with each other. We temporarily named the novel dependovirus California sea lion adeno-associated virus 1 (Csl-AAV1).

The genome organization of CsI-AAV1 was similar to those of other known AAVs, with two large ORFs (see Fig. S5A in the supplemental material) encoding the putative nonstructural (Rep) and capsid (VP) proteins, respectively. Multiple Rep and VP proteins may be produced by alternative initiation or mRNA splicing. The left ORF encoded a putative Rep1 protein of 600 aa, which showed 64% amino acid similarity to bovine and caprine AAVs. The putative VP1 protein consisted of 718 aa, showing the highest similarity (63% amino acid similarity) with AAV11 from cynomolgus monkey. Phylogenetic analyses of the VP1 protein revealed that CsI-AAV1 fell inside the mammalian AAV clade and was most closely related to the bovine AAV cluster (see Fig. S5B in the supplemental material).

DISCUSSION

This study describes the composition of the viral communities in the feces of three groups of California sea lions. Pups showed the smallest overall number of eukaryotic virus reads and the smallest ratio of eukaryotic virus/bacteriophage sequences, possibly reflecting protection by maternal antibodies in these approximately 3-month-old animals. Diseased animals in rehabilitation showed the largest overall number of eukaryotic virus reads and the largest ratio of eukaryotic virus/phage sequences, an observation that may reflect increased host susceptibility to enteric infections in these weakened animals.

Sixty-seven percent of phage-like reads in unweaned pups were related to siphoviruses, podoviruses, and myoviruses, similar to the phage compositions of human and equine feces reported by previous fecal viral metagenomic studies (8, 9, 11), showing that the majority of the phages (\sim 70%) were related to the tailed bacteriophages of the order *Caudovirales*. In contrast, the majority of the phage-like reads in juveniles (89%) and in diseased animals in rehabilitation (73%) were related to the single-stranded DNA microviruses. The less diverse phage composition in juveniles and animals in rehabilitation may reflect a similar change in their gut bacterial population.

All three animal groups showed a high level of coinfections, with averages of 1.6 distinct mammalian viruses for pups, 2.5 for juveniles, and 2.1 for animals in rehabilitation. This high level of coinfections with mammalian viruses may be an underestimate, and even deeper sequencing of the viral nucleic acids may have revealed an even greater diversity of viruses shed at lower levels. Whether the high number of viruses detected per animal is the result of frequent but rapidly resolving infections or of less frequent infections with long-term viral shedding, as seen, for example, with human bocavirus (40) will require analyses of longitudinally collected samples. Differences in viral load as reflected by the numbers of sequence reads might also reflect differences in the immune and health statuses of the hosts. The unweaned pups may have been partially protected from infections by maternal antibodies, whereas the animals in rehabilitation may have had reduced immune responses due to their concurrent diseases.

The presence of some eukaryotic viruses may also reflect the host diet. For example, feces from insectivorous bats contain a significant fraction of insect viruses and plant viruses (19, 35), and porcine circovirus and plant viruses can be detected in human feces (34, 67). Densoviruses, known to infect insects and crustaceans, were found exclusively in 6/14 sea lions in rehabilitation, which may be attributable to differences in the diet of captive versus wild sea lions, with the former being fed frozen herring (at times contaminated with nematode larvae and insects) rather than a variety of prey. A few other virus-like sequences of possible dietary origin were also detected in two animals in rehabilitation, namely, 2 nodavirus-related sequences (nodaviruses are known to infect insects and fishes) and 243 sequences which assembled into a >3-kb contig. This contig showed 23 and 34% amino acid identities with the capsid and nonstructural genes, respectively, of calhevirus, an unclassified member of the order Picornavirales thought to infect insects (28). The relative paucity of eukaryotic viruses from expected food sources may be due in part to the low number of genetically characterized fish and cephalopod viruses relative to insect or plant viruses, making their detection by sequence similarity more difficult. The still-nursing pups are also expected to be less exposed to viruses from solid foods.

We therefore detected viruses of mammalian origin belonging to multiple RNA and DNA virus families in California sea lion feces. Except for SMSV, none of these viruses had close homologues among previously described viruses. The sapelovirus, sapovirus, norovirus, bocavirus, and dependovirus sequences are the first ones from a marine mammal host to be reported.

Astroviruses had the highest prevalence and were found in 51% of the California sea lion fecal samples. A greater level of astrovirus genetic diversity than that previously reported (Csl AstV1 to Csl AstV3) (50) was noted here (Csl AstV4 to Csl AstV11) in this small population sampling. Bocaviruses

showed the second highest prevalence, with 38% of animals being infected with diverse variants. Overall, astroviruses and bocaviruses were the most abundant eukaryotic viruses detected, consisting of approximately 70% of the total eukaryotic virus reads. Whether these common viral infections are always commensal or can at times be pathogenic may depend on the overall health and immune status of the host and the presence of coinfections.

Our study provides an overview of the fecal virome of the California sea lion and significantly increases the diversity of viruses known to infect marine mammals, which now includes sapeloviruses, sapoviruses, noroviruses, bocavirus, and dependovirus. Characterization of the current baseline fecal virome of the California sea lion will help monitor future changes in its composition, which may be associated with disease outbreaks or population declines (22, 30, 57, 65).

ACKNOWLEDGMENTS

We acknowledge NHLBI grants R01HL083254 and R01HL105770, the BSRI for sustained support to E.D., and NSF award CNS-0619926 to the Bio-X2 cluster at Stanford University for computer resources.

We thank Jennifer Soper and Denise Creig from TMMC for sample collection and Robert DeLong for logistic support on San Miguel Island.

REFERENCES

- Allander, T., S. U. Emerson, R. E. Engle, R. H. Purcell, and J. Bukh. 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. Proc. Natl. Acad. Sci. U. S. A. 98:11609–11614.
- Allander, T., et al. 2005. Cloning of a human parvovirus by molecular screening of respiratory tract samples. Proc. Natl. Acad. Sci. U. S. A. 102:12891– 12896.
- Arthur, J. L., G. D. Higgins, G. P. Davidson, R. C. Givney, and R. M. Ratcliff. 2009. A novel bocavirus associated with acute gastroenteritis in Australian children. PLoS Pathog. 5:e1000391.
- Barrett, T., P. Wohlsein, C. A. Bidewell, and S. F. Rowell. 2004. Canine distemper virus in a Californian sea lion (Zalophus californianus). Vet. Rec. 154:334–336.
- Blasius, M. E., and G. D. Goodmanlowe. 2008. Contaminants still high in top-level carnivores in the Southern California Bight: levels of DDT and PCBs in resident and transient pinnipeds. Mar. Pollut. Bull. 56:1973–1982.
- Blomstrom, A. L., et al. 2009. Detection of a novel porcine boca-like virus in the background of porcine circovirus type 2 induced postweaning multisystemic wasting syndrome. Virus Res. 146:125–129.
- 7. Bonner, W. 2004. Seals and sea lions of the world. Facts on File, New York, NY.
- Breitbart, M., et al. 2008. Viral diversity and dynamics in an infant gut. Res. Microbiol. 159:367–373.
- Breitbart, M., et al. 2003. Metagenomic analyses of an uncultured viral community from human feces. J. Bacteriol. 185:6220–6223.
- Buckles, E. L., et al. 2006. Otarine herpesvirus-1, not papillomavirus, is associated with endemic tumours in California sea lions (Zalophus californianus). J. Comp. Pathol. 135:183–189.
- Cann, A. J., S. E. Fandrich, and S. Heaphy. 2005. Analysis of the virus population present in equine faeces indicates the presence of hundreds of uncharacterized virus genomes. Virus Genes 30:151–156.
- 12. Cheng, W. X., et al. 2010. Identification and nearly full-length genome characterization of novel porcine bocaviruses. PLoS One 5:e13583.
- Chevreux, B. 2005. MIRA: an automated genome and EST assembler. Ruprecht-Karls University, Heidelberg, Germany.
- Chiba, S., et al. 1979. An outbreak of gastroenteritis associated with calicivirus in an infant home. J. Med. Virol. 4:249–254.
- Colegrove, K. M., et al. 2010. Polyomavirus infection in a free-ranging California sea lion (Zalophus californianus) with intestinal T-cell lymphoma. J. Vet. Diagn. Invest. 22:628–632.
- Coria-Galindo, E., et al. 2009. Rotavirus infections in Galapagos sea lions. J. Wildl. Dis. 45:722–728.
- David, S. T., et al. 2007. An outbreak of norovirus caused by consumption of oysters from geographically dispersed harvest sites, British Columbia, Canada, 2004. Foodborne Pathog. Dis. 4:349–358.
- Dierauf, L. A., L. J. Lowenstine, and C. Jerome. 1981. Viral hepatitis (adenovirus) in a California sea lion. J. Am. Vet. Med. Assoc. 179:1194–1197.
- 19. Donaldson, E. F., et al. 2010. Metagenomic analysis of the viromes of three

North American bat species: viral diversity among different bat species that share a common habitat. J. Virol. **84**:13004–13018.

- Emerson, S. U., et al. 2005. Caliciviridae, p. 353–369. *In C. Fauquet, M. Mayo, J. Maniloff, U. Desselberger, and L. Ball (ed.), Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, CA.*
- Flotte, T. R., and K. I. Berns. 2005. Adeno-associated virus: a ubiquitous commensal of mammals. Hum. Gene Ther. 16:401–407.
- Grachev, M. A., et al. 1989. Distemper virus in Baikal seals. Nature 338:209– 210.
- Greig, D. J., F. M. D. Gulland, and C. Kreuder. 2005. A decade of live California sea lion (Zalophus californianus) strandings along the central California coast: causes and trends, 1991-2000. Aquat. Mammals 31:40–51.
- Guix, S., A. Bosch, and R. M. Pinto. 2005. Human astrovirus diagnosis and typing: current and future prospects. Lett. Appl. Microbiol. 41:103–105.
- Hernandez-Morga, J., J. Leon-Felix, F. Peraza-Garay, B. G. Gil-Salas, and C. Chaidez. 2009. Detection and characterization of hepatitis A virus and norovirus in estuarine water samples using ultrafiltration–RT-PCR integrated methods. J. Appl. Microbiol. 106:1579–1590.
- Kahn, J. 2008. Human bocavirus: clinical significance and implications. Curr. Opin. Pediatr. 20:62–66.
- Kapoor, A., et al. 2010. Identification and characterization of a new bocavirus species in gorillas. PLoS One 5:e11948.
- Kapoor, A., P. Simmonds, W. I. Lipkin, S. Zaidi, and E. Delwart. Use of nucleotide composition analysis to infer hosts for three novel picorna-like viruses. J. Virol. 84:10322–10328.
- Kapoor, A., et al. 2009. A newly identified bocavirus species in human stool. J. Infect. Dis. 199:196–200.
- Kennedy, S., et al. 2000. Mass die-off of Caspian seals caused by canine distemper virus. Emerg. Infect. Dis. 6:637–639.
- Kennedy-Stoskopf, S., M. K. Stoskopf, M. A. Eckhaus, and J. D. Strandberg. 1986. Isolation of a retrovirus and a herpesvirus from a captive California sea lion. J. Wildl. Dis. 22:156–164.
- King, D. P., et al. 2002. Otarine herpesvirus-1: a novel gammaherpesvirus associated with urogenital carcinoma in California sea lions (Zalophus californianus). Vet. Microbiol. 86:131–137.
- Kumar, S., M. Nei, J. Dudley, and K. Tamura. 2008. MEGA: a biologistcentric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. 9:299–306.
- Li, L., et al. 2010. Multiple diverse circoviruses infect farm animals and are commonly found in human and chimpanzee feces. J. Virol. 84:1674–1682.
- Li, L., et al. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. J. Virol. 84:6955–6965.
- Lipscomb, T. P., et al. 2000. Common metastatic carcinoma of California sea lions (Zalophus californianus): evidence of genital origin and association with novel gammaherpesvirus. Vet. Pathol. 37:609–617.
- Manteufel, J., and U. Truyen. 2008. Animal bocaviruses: a brief review. Intervirology 51:328–334.
- Martella, V., G. Elia, and C. Buonavoglia. 2008. Canine distemper virus. Vet. Clin. North Am. Small Anim. Pract. 38:787–797.
- Martella, V., et al. 2008. Identification of a porcine calicivirus related genetically to human sapoviruses. J. Clin. Microbiol. 46:1907–1913.
- Martin, E. T., et al. Frequent and prolonged shedding of bocavirus in young children attending daycare. J. Infect. Dis. 201:1625–1632.
- McClenahan, S. D., et al. 2008. Genomic characterization of novel marine vesiviruses from Steller sea lions (Eumetopias jubatus) from Alaska. Virus Res. 138:26–35.
- 42. Monroe, S., M. Carter, J. Hermann, D. Mitchell, and A. Sanchez-Fauquier. 2005. Astroviridae, p. 859–864. *In C. Fauquet, M. Mayo, J. Ma*niloff, U. Desselberger, and L. Ball (ed.), Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, CA.
- Neill, J. D., R. F. Meyer, and B. S. Seal. 1995. Genetic relatedness of the caliciviruses: San Miguel sea lion and vesicular exanthema of swine viruses constitute a single genotype within the Caliciviridae. J. Virol. 69:4484–4488.
- Newman, S. J., and S. A. Smith. 2006. Marine mammal neoplasia: a review. Vet. Pathol. 43:865–880.
- Ng, T. F., W. K. Suedmeyer, E. Wheeler, F. Gulland, and M. Breitbart. 2009. Novel anellovirus discovered from a mortality event of captive California sea lions. J. Gen. Virol. 90:1256–1261.
- Nollens, H. H., et al. 2006. Seroepidemiology of parapoxvirus infections in captive and free-ranging California sea lions Zalophus californianus. Dis. Aquat. Organ. 69:153–161.
- Nollens, H. H., et al. 2006. Pathology and preliminary characterization of a parapoxvirus isolated from a California sea lion (Zalophus californianus). J. Wildl. Dis. 42:23–32.
- Pomeroy, L. W., O. N. Bjornstad, and E. C. Holmes. 2008. The evolutionary and epidemiological dynamics of the paramyxoviridae. J. Mol. Evol. 66:98– 106.
- Ramig, R. F., M. Ciarlet, P. P. C. Mertens, and T. S. Dermody. 2005. Rotavirus, p. 859–864. In C. Fauquet, M. Mayo, J. Maniloff, U. Desselberger,

and L. Ball (ed.), Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, CA.

- Rivera, R., H. H. Nollens, S. Venn-Watson, F. M. Gulland, and J. F. Wellehan, Jr. 2010. Characterization of phylogenetically diverse astroviruses of marine mammals. J. Gen. Virol. 91:166–173.
- Saitoh, M., H. Kimura, K. Kozawa, O. Nishio, and A. Shoji. 2007. Detection and phylogenetic analysis of norovirus in Corbicula fluminea in a freshwater river in Japan. Microbiol. Immunol. 51:815–822.
- Sharp, C. P., et al. 2010. Widespread infection with homologues of human parvoviruses B19, PARV4, and human bocavirus of chimpanzees and gorillas in the wild. J. Virol. 84:10289–10296.
- 53. Smith, A. W., T. G. Akers, S. H. Madin, and N. A. Vedros. 1973. San Miguel sea lion virus isolation, preliminary characterization and relationship to vesicular exanthema of swine virus. Nature 244:108–110.
- Smith, A. W., et al. 1998. In vitro isolation and characterization of a calicivirus causing a vesicular disease of the hands and feet. Clin. Infect. Dis. 26:434–439.
- Smith, A. W., et al. 2006. Vesivirus viremia and seroprevalence in humans. J. Med. Virol. 78:693–701.
- Smith, A. W., D. E. Skilling, A. H. Dardiri, and A. B. Latham. 1980. Calicivirus pathogenic for swine: a new serotype isolated from opaleye Girella nigricans, an ocean fish. Science 209:940–941.
- Smith, A. W., N. A. Vedros, T. G. Akers, and W. G. Gilmartin. 1978. Hazards of disease transfer from marine mammals to land mammals: review and recent findings. J. Am. Vet. Med. Assoc. 173:1131–1133.
- Stanway, G., et al. 2005. Picornaviridae, p. 859–864. *In C. Fauquet, M. Mayo, J. Maniloff, U. Desselberger, and L. Ball (ed.), Virus taxonomy: eighth*

report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, CA.

- Tattersall, P., et al. 2005. Parvoviridae, p. 353–369. *In C. Fauquet, M. Mayo, J. Maniloff, U. Desselberger, and L. Ball (ed.), Virus taxonomy: eighth report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, CA.*
- 60. Terio, V., et al. 2010. Norovirus in retail shellfish. Food Microbiol. 27:29-32.
- Usuku, S., M. Kumazaki, K. Kitamura, O. Tochikubo, and Y. Noguchi. 2008. An outbreak of food-borne gastroenteritis due to sapovirus among junior high school students. Jpn. J. Infect. Dis. 61:438–441.
- Victoria, J. G., et al. 2009. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J. Virol. 83:4642–4651.
- Victoria, M., et al. 2010. Assessment of norovirus contamination in environmental samples from Florianopolis City, Southern Brazil. J. Appl. Microbiol. 109:231–238.
- Wellehan, J. J. 2010. Discovery, phylogenetic analysis, diagnostic test development, and surveillance of the astroviruses of marine mammals. Ph.D. thesis. University of Florida, Gainesville, FL.
- Wilder, F. W., and A. H. Dardiri. 1978. San Miguel sea lion virus fed to mink and pigs. Can. J. Comp. Med. 42:200–204.
- Zadori, Z., R. Stefancsik, T. Rauch, and J. Kisary. 1995. Analysis of the complete nucleotide sequences of goose and muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2. Virology 212:562–573.
- Zhang, T., et al. 2006. RNA viral community in human feces: prevalence of plant pathogenic viruses. PLoS Biol. 4:e3.