SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41564-018-0227-2

nature microbiology

In the format provided by the authors and unedited.

The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses

Tracey Goldstein[1,14*], Simon J. Anthony[2,3,4,14*], Aiah Gbakima[5], Brian H. Bird[1], James Bangura[5], Alexandre Tremeau-Bravard[1], Manjunatha N. Belaganahalli[1], Heather L. Wells[2], Jasjeet K. Dhanota[1], Eliza Liang[2,4], Michael Grodus[2], Rohit K. Jangra[6], Veronica A. DeJesus[6], Gorka Lasso[7], Brett R. Smith[1], Amara Jambai[8], Brima O. Kamara[9], Sorie Kamara[10], William Bangura[11], Corina Monagin[1,12], Sagi Shapira[7,13], Christine K. Johnson[1], Karen Saylors[12], Edward M. Rubin[12], Kartik Chandran[6], W. Ian Lipkin[2,3] and Jonna A. K. Mazet[1]

[1]One Health Institute & Karen C. Drayer Wildlife Health Center, School of Veterinary Medicine, University of California, Davis, CA, USA. [2]Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY, USA. [3]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA. [4]EcoHealth Alliance, New York, NY, USA. [5]Metabiota, Inc. Sierra Leone, Freetown, Sierra Leone. [6]Department of Microbiology and Immunology, Albert Einstein College of Medicine, New York, NY, USA. [7]Department of Systems Biology, Irving Cancer Research Center, Columbia University, New York, NY, USA. [8]Ministry of Health and Sanitation, Freetown, Sierra Leone. [9]Ministry of Health and Sanitation, Makeni, Sierra Leone. [10]Livestock and Veterinary Services Division, Ministry of Agriculture, Forestry and Food Security, Freetown, Sierra Leone. [11]Forestry and Wildlife Division, Ministry of Agriculture, Forestry and Food Security, Freetown, Sierra Leone. [12]Metabiota, Inc., San Francisco, CA, USA. [13]Department of Microbiology & Immunology, Columbia University, New York, NY, USA. [14]These authors contributed equally: Tracey Goldstein, Simon J. Anthony. *e-mail: tgoldstein@ucdavis.edu; sja2127@cumc.columbia.edu

Supplementary Information

The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses

Supplementary Table 1. Summary of bats tested for Filoviruses and BOMV.

Bat Species	Bat Family	No. Tested	No. Positive by Filovirus cPCR	No. Positive by Ebolavirus genus PCR	No. Positive by qRT- PCR for BOMV
Insectivorous Bats					
Chaerephon pumilus	Molossidae	55	3	2	4
Glauconycteris poensis	Vespertillionidae	1			
Hipposideros abae	Hipposiderdae	7			
Hipposideros jonesi	Hipposiderdae	4			
Hipposideros ruber	Hipposiderdae	50			
Mops condylurus	Molossidae	52	1		1
Myotis bocagii	Vespertillionidae	3			
Neoromicia rendalli	Vespertillionidae	2			
Nycteris hispida	Nycteridae	1			
Pipistrellus nanulus	Vespertillionidae	3			
Rhinolophus fumigatus	Rhinolophidae	3			
Rhinolophus landeri	Rhinolophidae	1			
Rhinopoma microphyllum	Rhinopomatidae	1			
Scotophilus viridis	Vespertillionidae	26			
Unidentified Chaerephon bat	Molossidae	1			
Unidentified Molossid bat	Molossidae	1			
Unidentified Hipposiderod bat	Hipposiderdae	1			
Unidentified Nycterid bat	Nycteridae	3			
Fruit Bats					
Eidolon helvum	Pteropodidae	2			
Epomophorus gambianus	Pteropodidae	1			
Epomops buettikoferi	Pteropodidae	2			
Micropteropus pusillus	Pteropodidae	2			
Myonycteris angolensis	Pteropodidae	14			
Unidentified Myonycteris/Epomophorous bat	Pteropodidae	3			
Unidentified Epomophorous bat	Pteropodidae	5			
Total		**244**	**4**	**2**	**5**

Supplementary Table 2. Summary of sites, attributes and animals sampled

Site No*	District Location	Urban/Rural	Species sampled (No animals†)	Sampling Month	Habitat Type
1	Western Area Urban	Urban	*Canis lupus familiaris* (20)	July 2016	In and around dwellings
2	Western Area Urban	Peri-urban	*Canis lupus familiaris* (21)	July 2016	In and around dwellings
3	Western Area Urban	Urban	*Canis lupus familiaris* (9)	July 2016	In and around buildings and dwellings
4	Western Area Urban	Urban	*Canis lupus familiaris* (4)	July 2016	In and around buildings
5	Kambia	Rural	*Mastomys erythroleucus* (2) *Mus musculus* (10) *Rattus rattus* (5)	Sept 2016	In and around dwellings, Natural areas
6	Kambia	Rural	*Canis lupus familiaris* (9) *Felis catus* (1) *Mastomys erythroleucus* (1) *Mus musculus* (2) Rattus rattus (2)	Sept 2016	In and around dwellings, Natural areas
7	Bombali	Rural	*Canis lupus familiaris* (29) *Felis catus* (2) *Nycteris hispida* (1) *Rhinolophus fumigatus* (1) *Rhinolophus landeri* (1) *Chaerephon pumilus* (1†/2) *Mops condylurus* (3) *Hipposideros abae* (7) *Hipposideros jonesi* (2) *Hipposideros ruber* (34) Unidentified Hipposiderod bat (1)	March 2016 May 2016 July 2016	In and around dwellings, Natural Areas
8	Bombali	Urban	*Canis lupus familiaris* (14) *Rhinolophus fumigatus* (1) *Scotophilus viridis* (8) *Chaerephon pumilus* (1†/5) Unidentified Nycterid bat (2) *Epomophorus gambianus* (1) Unidentified Myonycteris/Epomorphorous bat (1)	June 2016	In and around dwellings
9	Bombali	Rural	*Canis lupus familiaris* (8) *Pipistrellus nannulus* (1) *Chaerephon pumilus* (2) *Epomops buettikoferi* (2) *Myonycteris angolensis* (1) Unidentified Epomophorus bat (1)	June 2016	In and around dwellings, Crop fields, Natural areas
10	Bombali	Peri-urban	*Canis lupus familiaris* (35) *Scotophilus viridis* (17) *Mops condylurus* (9) *Chaerephon pumilus* (2) *Micropteropus pusillus* (1) Unidentified Epomophorus bat (3)	May 2016	In and around dwellings, Crop fields, Natural areas
11	Bombali	Rural	*Canis lupus familiaris* (15) *Rattus rattus* (2) *Myotis bocagii* (1) *Neoromicia rendalli* (2) *Pipistrellus nannulus* (1) *Mops condylurus* (1†/21) *Hipposideros jonesi* (1) *Eidolon helvum* (2) *Myonycteris angolensis* (1) Unidentified Epomophorus bat (1)	March 2016 May 2016	In and around dwellings

12	Bombali	Rural	Rattus rattus (2) *Mops condylurus* (19)	March 2016	In and around dwellings, Crop fields
13	Bombali	Rural	*Canis lupus familiaris* (30) *Chaerephon pumilus* (5)	June 2016	Natural areas
14	Bombali	Rural	*Canis lupus familiaris* (19) *Myotis bocagii* (2) *Pipistrellus nannulus* (1) *Scotophilus viridis* (1) *Chaerephon pumilus* (2†/18) *Glauconycteris poensis* (1) *Hipposideros ruber* (2) Unidentified Nycterid bat (1) *Micropteropus pusillus* (1)	May 2016	In and around dwellings, Natural areas
15	Bombali	Rural	*Canis lupus familiaris* (4)	May 2016	In and around dwellings
16	Bombali	Rural	*Canis lupus familiaris* (5) *Rhinopoma microphyllum* (1) *Hipposideros jonesi* (1) *Hipposideros ruber* (14) *Rhinolophus fumigatus* (1) *Chaerephon pumilus* (21) Unidentified Chaerephon bat (1) Unidentified Molossid bat (1) *Myonycteris angolensis* (5)	June 2016	In and around dwellings, Natural areas
17	Bombali	Rural	*Canis lupus familiaris* (5) *Myonycteris angolensis* (7) Unidentified Myonycteris/Epomorphorous bat (2)	July 2016	In and around dwellings, Natural areas
18	Koinadugu	Rural	*Mastomys erythroleucus* (1)	Sept 2016	In and around dwellings, Natural areas
19	Kono	Rural	*Canis lupus familiaris* (6) *Mastomys natalensis* (10) *Rattus rattus* (2)	Sept 2016	In and around dwellings, Crop fields
20	Kono	Rural	*Canis lupus familiaris* (7) *Felis catus* (2) *Mastomys natalensis* (1) *Mastomys erythroleucus* (1) *Rattus rattus* (5)	Sept 2016	In and around dwellings

* Number corresponds with numbers on map in Figure 1

† Number of animals testing positive for BOMV

Supplementary Table 3. Amino acids indicated to be under positive selection for at least 2 of the 4 datamonkey algorithms tested: SLAC, FEL, MEME, and FUBAR.

Gene	Position
NP	**3(4)**, 11(2), 108(2), 502(2), 553(2), 577(2), 627(2)
VP35	63(2)
VP40	67(2)
$GP_{1,2}$	**310(4), 318(3), 321(3), 332(3)**
VP30	276(2)
VP24	None
L	202(2), 1661(2), 1731(2), 1733(2), 1737(2), 1752(2), 1774(2), 2171(2)

Numbering is according to EBOV.

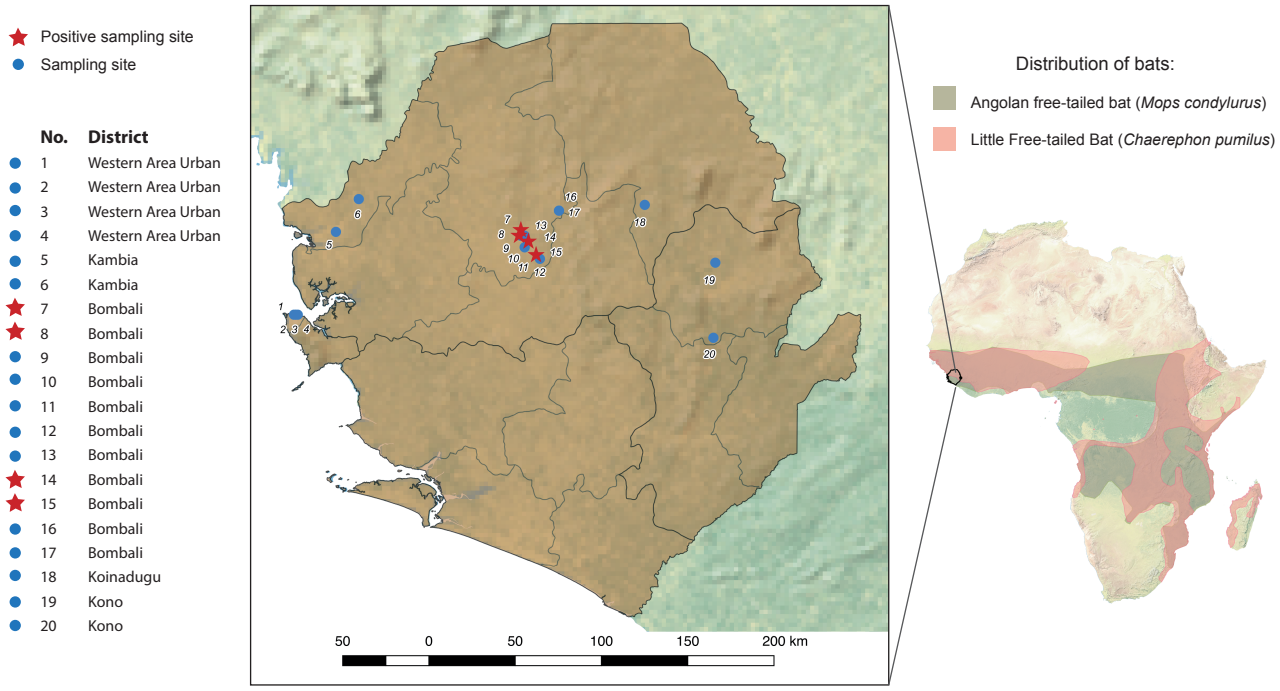
Number in parentheses is the number of tests that indicated positive selection.

Sites indicated by 3 or more methods are bolded

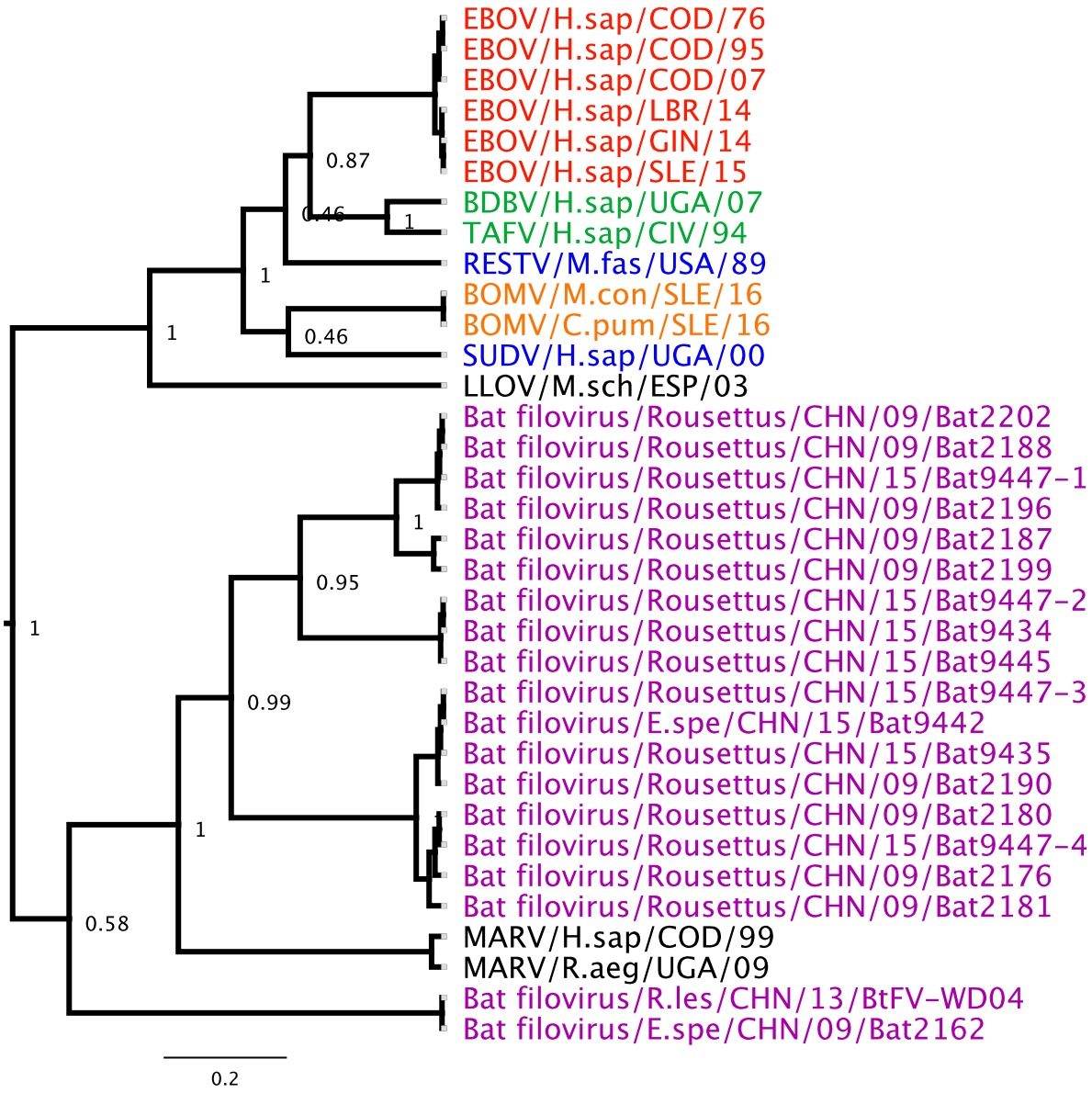
Supplementary Table 4. Comparison on reactivity of GP_1 peptides and recombinant EBOV $GP_{1,2}$ with polyclonal antibodies against EBOV, SUDV, BDBV and TAFV (++++ = O.D. > 4.0, +++ = O.D. 3.0-4.0, ++ = O.D. 2.0-3.0, + = O.D. >0.5-2.0). Greatest cross reaction was observed with peptide GP-100 which showed the highest sequence similarity to known ebolaviruses; and no reactivity was observed with GP-471 (lowest sequence similarity). GP-313 and GP-378 only bound to anti- EBOV and TAFV antibodies respectively, as expected.

	Antigen					
Antibody	**EBOV Recombinant $GP_{1,2}$**	**GP-100 (BOMV)**	**GP-270 (BOMV)**	**GP-471 (BOMV)**	**GP-313 (EBOV)**	**GP-378 (TAFV)**
Anti-EBOV	**++++**	**+++**	**+++**		**++**	
Anti-SUDV	**+++**	**+**				
Anti-BDBV	**++++**	**+**	**+**			
Anti-TAFV	**++**					**++**
Anti-RESTV	**+++**					

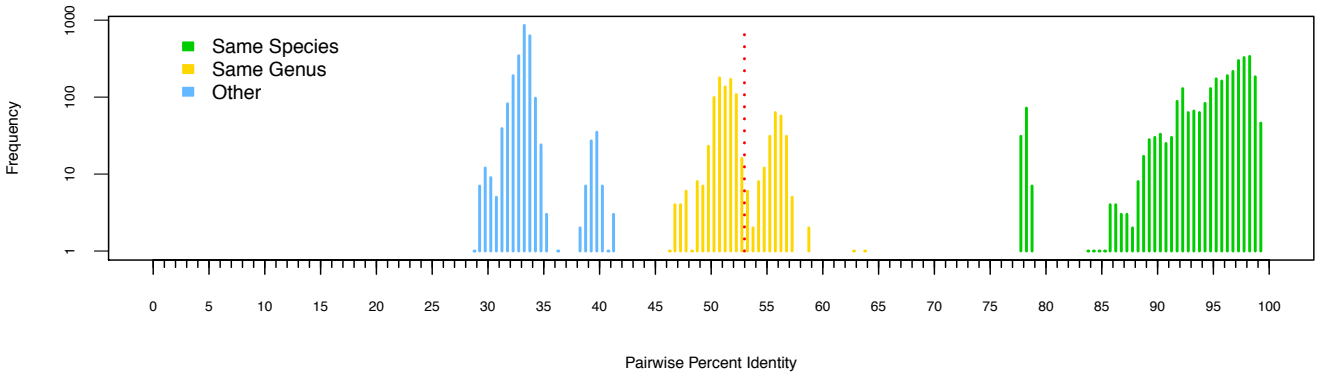
Supplementary Figure 1. Map showing the distribution of Angolan free-tailed bats (*Mops condylurus*) and Little free-tailed bats (*Chaerephon pumilus*) (based on International Union for Conservation of Nature [IUCN] data) and the animal sampling locations in Sierra Leone.



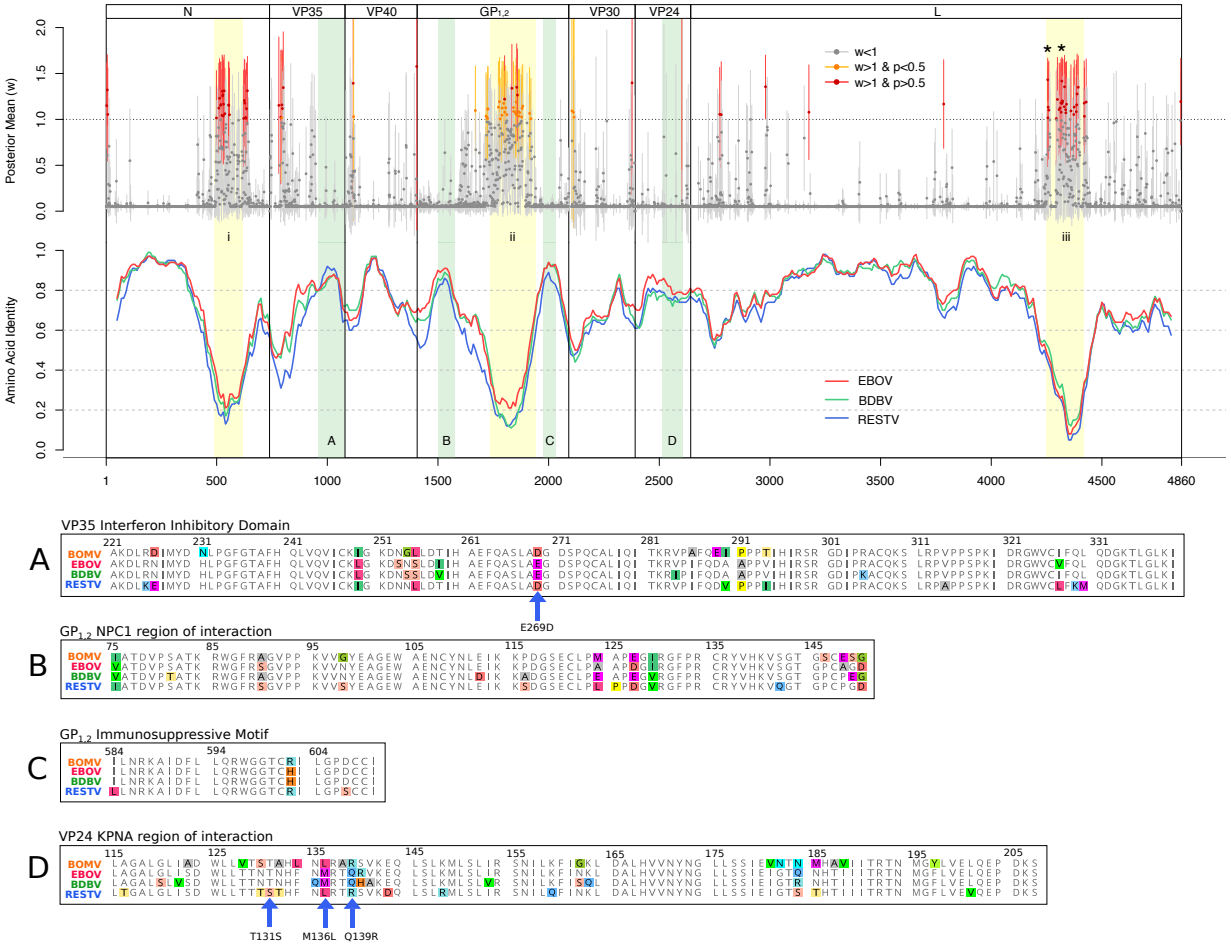
Supplementary Figure 2. Phylogenetic tree comparing the relationship of BOMV to other unclassified filoviruses from China (shown in purple) reported by Yang et al[27]. Trees were created using the 310bp nucleotide sequences from the Chinese filoviruses and the corresponding regions from complete genome filoviruses. Branch support for n = 34 filovirus sequences indicated by Bayesian posterior values derived from 1,000,000 generations. Genbank accession numbers used in phylogenetic and evolutionary analyses for Figure 1 and Supplementary Figure 2 are provided in Data Availability.



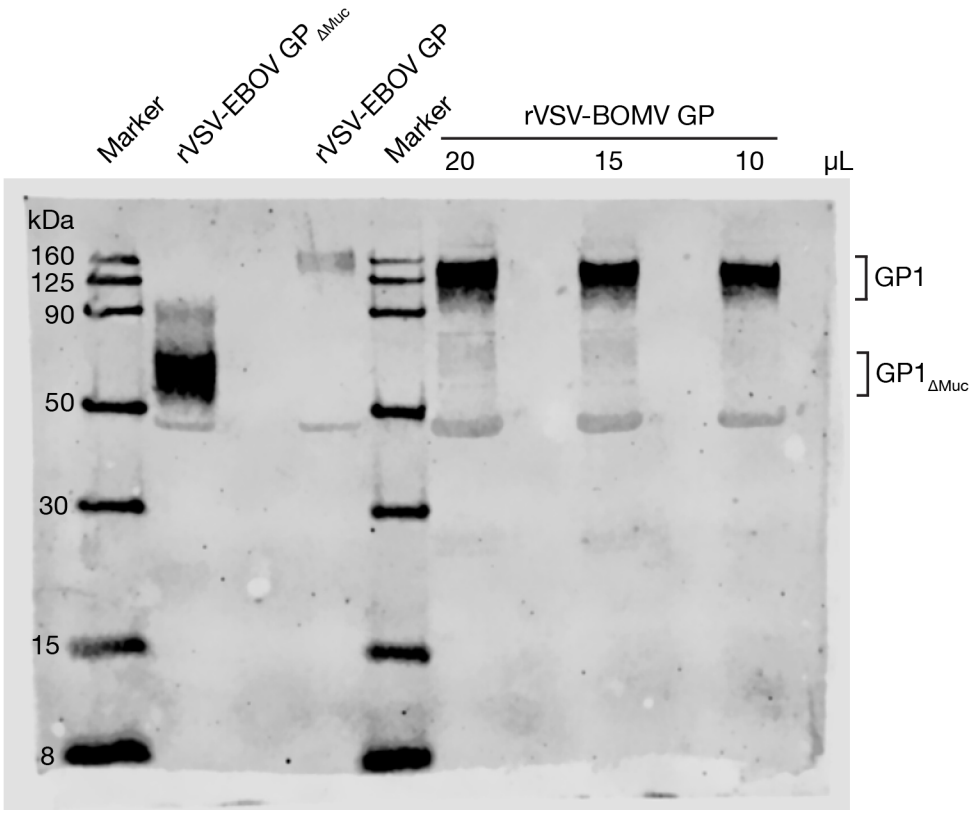
Supplementary Figure 3. A histogram of pairwise sequence identities between filovirus genomes generated by the PASC tool from NCBI. Sequences that were >99.5% identical were removed, leaving 112 non-redundant genomes. Bars are color-coded to indicate whether pairwise identities arose from two sequences classified as the same species, the same genus, or neither the same species or genus (e.g. EBOV and MARV). The maximum pairwise identity of BOMV (dashed red line) to other ebolaviruses was 52.98%, therefore satisfying requirements for the establishment of a novel species.



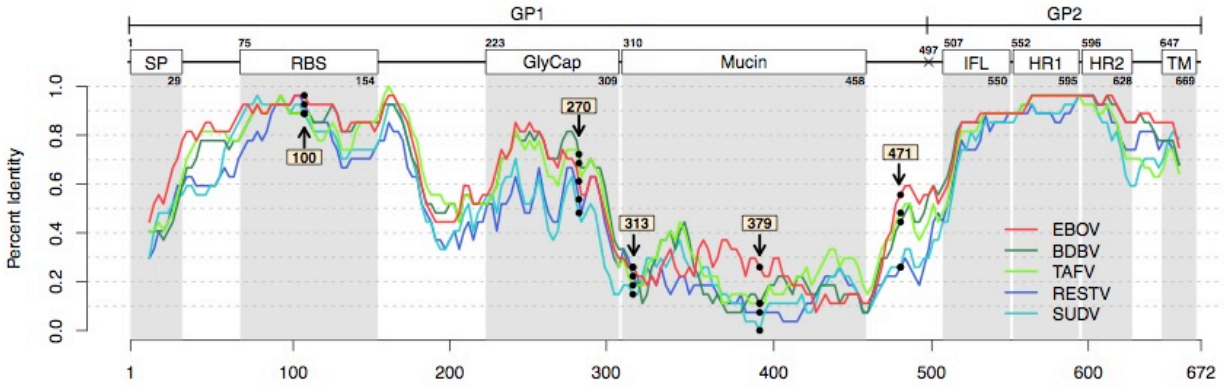
Supplementary Figure 4. Comparison of amino acid motifs of BOMV and three ebolaviruses showing variation in selective pressure and amino acid sequence identity across the BOMV genome. The nucleotide alignment was analyzed for evolutionary selection using the M7 and M8 codon models in the codeml package of PAML. Each position has an estimate of selection pressure w (the posterior mean *w*), a credible interval for the estimated posterior mean *w*, and a posterior probability *p* that the true value w > 1 and that the corresponding site is under positive selection pressure. The Bayes Empirical Bayes (BEB) posterior mean *w* ± the credible interval at each position is shown (top). Positions where the posterior mean *w* > 1 are colored orange if the corresponding posterior probability *p* < 0.5 and colored red if the posterior probability *p* > 0.5, indicating slight and moderate evidence for positive selection, respectively. Positions with *p* > 0.9 are marked with an asterisk, indicating strong evidence for positive selection at that position. The amino acid similarity (bottom) shows the percent sequence identity of EBOV (KC242801), BDBV (KU182911), and RESTV (NC_004161) ebolaviruses compared with the reference Bombali ebolavirus /Mops condylurus/ SLE/2016/PREDICT_SLAB00156 in a sliding window size of 100 with step size 10. Sections highlighted in green (A, B, C, D) are regions of high amino acid conservation and purifying selection. The amino acid sequences of these regions are shown below with amino acid changes highlighted. All amino acid numbering is according to EBOV. Blue arrows indicate Specificity Determining Positions identified by Pappalardo et al.[16], showing that some residues are more similar to EBOV, while others are more similar to RESTV. Sections highlighted in yellow (i, ii, iii) are regions of high amino acid variation that show enriched positive selection.



Supplementary Figure 5. Raw data showing BOMV $GP_{1,2}$ incorporation into rVSV particles. Dilutions of rVSV particles were resolved on an SDS-PAGE and incorporation of $GP_{1,2}$ was detected by immunoblotting using ebolavirus anti-GP_1 serum. rVSVs carrying full length EBOV GP or EBOV GPΔMuc lacking the heavily glycosylated mucin domain (corresponding to amino acids 309-489) were used as positive controls. Mucin residues 309-489 are dispensable for viral entry in tissue culture.



Supplementary Figure 6. Similarity plot showing sequence identity (amino acid) across the BOMV glycoprotein compared with the five known ebolaviruses (total sequences n = 6). All sequences compared independently to BOMV. Highest conservation is observed in the receptor binding site (RBS) and highest variability in the mucin-like domain. Peptides used in the ELISA are indicated, and numbered based on the position of the first amino acid residue. Sequences of each peptide are also provided below, together with their amino acid percent identity to the other ebolaviruses. As shown, Peptide-100 is the most conserved and would be expected to cross react with antibodies from all ebolaviruses, whereas Peptides -313 and -379 are found in highly variable regions and are expected to be specific for EBOV and TAFV respectively.



Peptide-100

BOMV	EAGEWAENCY NLEIKKPDGS ECLP	
EBOV	EAGEWAENCY NLEIKKPDGS ECLP	100%
BDBV	EAGEWAENCY NLDIKKADGS ECLP	91.7%
TAFV	EAGEWAENCY NLAIKKVDGS ECLP	91.7%
RESTV	EAGEWAENCY NLEIKKSDGS ECLP	95.8%
SUDV	EAGEWAENCY NLEIKKPDGS ECLP	100%

Peptide-270

BOMV	IWQVNSRVDT DVGEWAFWEN KKNLKKSFP	
EBOV	IWKVNPEIDT TIGEWAFWET KKNLTRKIR	58.6%
BDBV	IWKVNPTVDT GVGEWAFWEN KKNFTKTLS	69.0%
TAFV	IWKINPTVDT SMGEWAFWEN KKNFKKTLS	65.5%
RESTV	TWTLDPKIEP DVGEWAFWET KKNFSQQLH	44.8%
SUDV	IWTLDANINA DIGEWAFWEN KKNLSEQLR	51.7%

Peptide-471

BOMV	SIDGPGLLTN TLTGVARLIT GAGRAKR	
EBOV	SSGKLGLITN TIAGVAGLIT GGRRTRR	55.6%
BDBV	ESTEPGPLTN TTRGAANLLT GSRRTRR	48.1%
TAFV	EISGPGFLTN TIRGVTNLLT GSRRKRR	51.9%
RESTV	GPSQPGLTIN TVSKVADSLS PTRKQKR	33.3%
SUDV	SN---GLITS TVTGILGSLG LRKRSRR	33.3%

Peptide-313 (EBOV-specific)

EBOV	NGAKNISGQS PARTSSDPGT NTTTEDHKIM AS	
BOMV	RAADSEHDAH PPEHTPGPDS NPTINDNTEL VT	18.8%
BDBV	RAQDPGSNQK TKVTPTSFAN NQTSKNHEDL VP	18.8%
TAFV	ETQNQVLDTT ATVSPPISAH NHAAEDHKEL VS	18.8%
RESTV	THTNNSSDQS PAGTVQGKIS YHPPANNSEL VP	21.9%
SUDV	LNETEDDDAA SSRITKGRIS DRATRKYSDL VP	6.3%

Peptide-379 (TAFV-specific)

TAFV	INARNTDHTK SFIGLEGPQE DHSTTQPAKT TSQ	
BOMV	PPAAQHDN-- PRNSPTPPSP IESDITGSTQ ALD	12.1%
EBOV	SPQSLTTKPG PDNSTHNTPV YKLDISEATQ VEQ	9.1%
BDBV	MEEQTTSHYE PPNISRNHQE RNNTAHPETL ANN	18.2%
RESTV	TANPMTTTIA PSP-TMTSEV DNNVPSEQPN NTA	6.1%
SUDV	TQLTITETAA TIIGTNGNHM QISTIGIRPS SSQ	24.2%