HUMAN EBOLA VIRUS SPECIES AND COMPOSITIONS AND METHODS THEREOF

Inventors: Jonathan S. Towner, Atlanta, GA (US); Stuart T. Nichol, Atlanta, GA (US); James A. Comer, Atlanta, GA (US); Thomas G. Ksiazek, Atlanta, GA (US); Pierre E. Rollin, Atlanta, GA (US)

Assignee: The Government of the US as Represented by the Secretary of the Dept. of health, Atlanta, GA (US)

Publication Classification

- Int. Cl.
  - A61K 35/76 (2006.01)
  - C07H 21/04 (2006.01)
  - C12N 7/04 (2006.01)
  - C07K 14/08 (2006.01)
  - A61K 38/02 (2006.01)
  - A61K 31/7088 (2006.01)
  - C07K 7/06 (2006.01)
  - C07K 7/08 (2006.01)
  - C12N 7/00 (2006.01)
  - C07H 21/02 (2006.01)

U.S. Cl. 424/93.6; 435/235.1; 536/23.72; 435/236; 530/350; 514/1.1; 514/44 R; 530/330; 530/329; 530/328; 530/327; 530/326; 530/325; 530/324

Abstrac

Compositions and methods including and related to the Ebola Bundibugyo virus (EboBun) are provided. Compositions are provided that are operable as immunogens to elicit and immune response or protection from EboBun challenge in a subject such as a primate. Inventive methods are directed to detection and treatment of EboBun infection.
Fig. 1
<table>
<thead>
<tr>
<th>Ebola Beauchamp '07</th>
<th>Ebola Zaire '96</th>
<th>Ebola Zaire '96</th>
<th>Ebola Zaire '76</th>
<th>Ebola Zaire '76</th>
<th>Ebola Zaire '76</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>GAGCAGACAGGAGCACTG</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>CCGGCTGGCTGGCTGGCT</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>TACGTTACCGGTTACGTC</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Gene</td>
<td>Sequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Bundibugyo '07</strong></td>
<td>TTATATATCTTCTAATAAAATACAGTATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola IC '94</strong></td>
<td>CAAAAGCCTACCAAAAAACAGGGTATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Zaire '76</strong></td>
<td>TTATATATCTTCTAATAAAATACAGTATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Bundibugyo '07</strong></td>
<td>GATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola IC '94</strong></td>
<td>GATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Zaire '76</strong></td>
<td>GATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Bundibugyo '07</strong></td>
<td>AATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola IC '94</strong></td>
<td>AATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Zaire '76</strong></td>
<td>AATTACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Bundibugyo '07</strong></td>
<td>ATCAATACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola IC '94</strong></td>
<td>ATCAATACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ebola Zaire '76</strong></td>
<td>ATCAATACATACCCAATGGATGAGGATTAAAGAGCATCCTTTATCTATGAAGGACAGTAAATTCTTTATATCTACCACTAATCTTATAGATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8910</td>
<td>8920</td>
<td>8930</td>
<td>8940</td>
<td>8950</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8910</td>
<td>8920</td>
<td>8930</td>
<td>8940</td>
<td>8950</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9010</td>
<td>9020</td>
<td>9030</td>
<td>9040</td>
<td>9050</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9110</td>
<td>9120</td>
<td>9130</td>
<td>9140</td>
<td>9150</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9210</td>
<td>9220</td>
<td>9230</td>
<td>9240</td>
<td>9250</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9310</td>
<td>9320</td>
<td>9330</td>
<td>9340</td>
<td>9350</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9410</td>
<td>9420</td>
<td>9430</td>
<td>9440</td>
<td>9450</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9510</td>
<td>9520</td>
<td>9530</td>
<td>9540</td>
<td>9550</td>
</tr>
<tr>
<td>Ebola Bundibugyo '07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola IC '94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ebola Zaïre '76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 2**
<table>
<thead>
<tr>
<th></th>
<th>Sample Type</th>
<th>Sequence Details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ebola Bundibugyo '07</td>
<td>AACAAGACGCATACATTAGGGATTTGGTTCATACATCTTCCTGCGCAGCACTACAGCGAGGTTACCTGGAACCTGTAATGAGATATTTAGC</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>AGGCAACAAAAAATCTGATATCTCAGCAGCAACGTCGCTGAGGAAGAAAACATCGCTAAGCTGAGGAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola IC '94</td>
<td>AGTTGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>AGGCAACAAAAAATCTGATATCTCAGCAGCAACGTCGCTGAGGAAGAAAACATCGCTAAGCTGAGGAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Bundibugyo '07</td>
<td>AGGCAACAAAAAATCTGATATCTCAGCAGCAACGTCGCTGAGGAAGAAAACATCGCTAAGCTGAGGAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>AGGCAACAAAAAATCTGATATCTCAGCAGCAACGTCGCTGAGGAAGAAAACATCGCTAAGCTGAGGAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola IC '94</td>
<td>AGTTGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>AGGCAACAAAAAATCTGATATCTCAGCAGCAACGTCGCTGAGGAAGAAAACATCGCTAAGCTGAGGAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Bundibugyo '07</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola IC '94</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Bundibugyo '07</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola IC '94</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>Ebola Zaire '76</td>
<td>GTGGACAGAAGGATCATCATACTCAAGGCTAATAGCCTGCGGCTGCTGAGGTC-CAACAGGGCCCCAGCTGACATCTGGGGAAGCAGGATACCACTGTTATCGCCAAGAAAAA</td>
</tr>
</tbody>
</table>

FIG. 2
HUMAN EBOLA VIRUS SPECIES AND COMPOSITIONS AND METHODS THEREOF

RELATED APPLICATIONS

This application claims priority benefit of U.S. Provisional Application 61/108,175 filed 24 Oct. 2008; the contents of which are hereby incorporated by reference.

DEPOSIT STATEMENT

The invention provides the isolated human Ebola (hEbola) viruses denoted as Bundibugyo (EboBun) deposited with the Centers for Disease Control and Prevention ("CDC"); Atlanta, Ga., United States of America) on Nov. 26, 2007 and accorded an accession number 200706291. This deposit was not made to an International Depository Authority (IDA) as established under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure, and is accepted by the Budapest Treaty deposit. The deposited organism is not acceptable by American Type Culture Collection (ATCC), Manassas, Va., an International Depository Authority (IDA) as established under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. Samples of the stated Deposit Accession No. 200706291 will be made available to approved facilities for thirty years from the date of deposit, and for the lifetime of the patent issuing from, or claiming priority to this application.

FIELD OF THE INVENTION

The invention is related to compositions and methods directed to a novel species of human Ebola (hEbola) virus.

BACKGROUND OF THE INVENTION

The family Filoviridae consists of two genera, Marburgvirus and Ebola virus, which have likely evolved from a common ancestor. The genus Ebola virus includes four species: Zaire, Sudan, Reston and Cote d'Ivoire (Ivory Coast) ebolaviruses, which have, with the exception of Reston and Cote d'Ivoire ebolaviruses, been associated with large hemorrhagic fever (HF) outbreaks in Africa with high case fatality (53-90%) .

Viruses of each species have genomes that are at least 30-40% divergent from one another, a level of diversity that presumably reflects differences in the ecological niche they occupy and in their evolutionary history. Identification of the natural reservoir of ebolaviruses remains somewhat elusive, although recent PCR and antibody data suggest that three species of arboreal fruit bats may be carriers of Zaire ebolavirus . No data has yet been published to suggest reservoirs for the Sudan, Reston and Cote d'Ivoire ebolavirus species. However, a cave-dwelling fruit bat has been recently implicated as a natural host for marburgvirus , supporting the hypothesis that different bat species may be the reservoir hosts for the various filoviruses.

Filovirus outbreaks are sporadic, sometimes interspersed by years or even decades of no apparent disease activity. The last new species of ebolavirus was discovered 14 years ago (1994), in Cote d'Ivoire (Ivory Coast), and involved a single non-fatal case, a veterinarian who performed an autopsy on an infected chimpanzee found in the Tai Forest . No further disease reports have been associated with Cote d'Ivoire ebolavirus, in contrast to Zaire and Sudan ebolaviruses which have each caused multiple large outbreaks over the same time period.

In late November 2007, HF cases were reported in the townships of Bundibugyo and Kikuyu in Bundibugyo District, Western Uganda. The outbreak continued through January 2008, and resulted in approximately 149 cases and 37 deaths . Laboratory investigation of the initial 29 suspect-case blood specimens by classic methods (antigen capture, IgM and IgG ELISA) and a recently developed random-primer pyrosequencing approach identified this to be an Ebola HF outbreak associated with a new discovered ebolavirus species. These specimens were negative when initially tested with highly sensitive real-time RT-PCR assays specific for all known Zaire and Sudan ebolaviruses and Marburg viruses. This new species is referred to herein as "the Bundibugyo species", abbreviated "EboBun".

Accordingly, compositions and methods directed to the new Ebola virus species are described herein and the most closely related Ebola Ivory Coast species, which compositions and methods are useful for diagnosis and prevention of human Ebola virus infection; including related vaccine development, and prevention of hemorrhagic fever in a human population.

SUMMARY OF THE INVENTION

The present invention is based upon the isolation and identification of a new human Ebola virus species, EboBun. EboBun was isolated from the patients suffering from hemorrhagic fever in a recent outbreak in Uganda. The isolated virus is a member of the Filoviridae family, a family of negative sense RNA viruses. Accordingly, the invention relates to the isolated EboBun virus that morphologically and phylogenetically relates to known members filoviridae.

In one aspect, the invention provides the isolated EboBun virus deposited with the Centers for Disease Control and Prevention ("CDC"); Atlanta, Ga., United States of America) on Nov. 26, 2007 and accorded an accession number 200706291, as stated in the paragraph entitled "DEPOSIT STATEMENT" supra.

In another aspect, the invention provides an isolated hEbola EboBun virus comprising a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of: a) a nucleotide sequence set forth in SEQ ID NO: 1; b) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 1 under stringent conditions; and c) a nucleotide sequence that has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to the SEQ ID NO: 1. In another aspect, the invention provides the complete genomic sequence of the hEbola virus EboBun.

In a related aspect, the invention provides nucleic acid molecules isolated from EboBun, or fragments thereof.

In another aspect, the invention provides proteins or polypeptides that are isolated from the EboBun, including viral proteins isolated from cells infected with the virus but not present in comparable uninfected cells; or fragments thereof. In one embodiment of the present invention, the amino acid sequences of the proteins or polypeptides are set forth in SEQ ID NOS: 2-9 and 59, or fragments thereof.

In a related aspect, the invention provides an isolated polypeptide encoded by the nucleic acid molecule of the inventive hEbola EboBun (Sequence ID No. 10) virus described above.
In another aspect, the invention provides an isolated hEbola EbolC virus comprising a nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of: a) a nucleotide sequence set forth in SEQ ID NO: 10; b) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 10 under stringent conditions; and c) a nucleotide sequence that has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to the SEQ ID NO: 10. In another aspect, the invention provides the complete genomic sequence of the hEbola virus EbolC.

In a related aspect, the invention provides nucleic acid molecules isolated from EbolC, or fragments thereof.

In another aspect, the invention provides proteins or polypeptides that are isolated from the EbolC, including viral proteins isolated from cells infected with the virus but not present in comparable uninfected cells; or fragments thereof. In one embodiment of the present invention, the amino acid sequences of the proteins or polypeptides are set forth in SEQ ID NOs: 11-19, or fragments thereof.

In a related aspect, the invention provides an isolated polypeptide encoded by the nucleic acid molecule of the inventive hEbola EbolC virus described above.

In other aspects, the invention relates to the use of the isolated hEbola virus for diagnostic and therapeutic methods based on EbBun, EbolC, or a combination thereof. In one embodiment, the invention provides a method of detecting in a biological sample an antibody immunospecific for the genus of West African Ebola Species constituting hEbola EbBun and EbolC virus using at least one of the inventive isolated hEbola virus described herein, or any of the inventive proteins or polypeptides described herein. In another specific embodiment, the invention provides a method of screening for an antibody which immunospecifically binds and neutralizes hEbola EbolBun. Such an antibody is useful for a passive immunization or immunotherapy of a subject infected with hEbola.

In another aspect, the invention provides an isolated antibody or an antigen-binding fragment thereof which immunospecifically binds to the hEbola virus of the invention described above.

In other aspects, the invention provides methods for detecting the presence, activity or expression of the Glade of Bundibungyo-Ivory Coast hEbola virus in a biological material, such as cells, blood, saliva, urine, feces and so forth; and specifically at least one of EbBun or EbolC.

In a related aspect, the invention provides a method for detecting the presence of the inventive hEbola virus described above in a biological sample, the method includes (a) contacting the sample with an agent that selectively binds to a West African hEbola virus; and (b) detecting whether the compound binds to the West African hEbola virus in the sample.

In another aspect, the invention provides a method for detecting the presence of the inventive polypeptide described above, in a biological sample, said method includes (a) contacting the biological sample with an agent that selectively binds to the polypeptide; and (b) detecting whether the agent binds to the polypeptide in the sample. In another aspect, the invention provides a method for detecting the presence of a first nucleic acid molecule derived from the inventive hEbola virus described above in a biological sample, the method comprising: (a) contacting the biological sample with an agent that selectively binds to the polypeptide; and (b) detecting whether the agent binds to the polypeptide in the sample.

In another aspect, the invention provides a method for propagating the hEbola virus in host cells comprising infecting the host cells with the inventive isolated hEbola virus described above, culturing the host cells to allow the virus to multiply, and harvesting the resulting virions. Also provided by the present invention are host cells infected with the inventive hEbola virus described above.

In another aspect, the invention provides a method of detecting in a biological sample the presence of an antibody that immunospecifically binds hEbola virus, the method comprising: (a) contacting the biological sample with the inventive host cell host described above; and (b) detecting the antibody bound to the cell.

In another aspect, the invention provides vaccine preparations, comprising the inventive hEbola virus, including recombinant and chimeric forms of the virus, nucleic acid molecules comprised by the virus, or protein subunits of the virus. The invention also provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of the inventive hEbola virus described above, and a pharmaceutically acceptable carrier. In one embodiment, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a protein extract of the inventive hEbola virus described above, or a subunit thereof; and a pharmaceutically acceptable carrier. In another embodiment, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 or a complement thereof, and a pharmaceutically acceptable carrier. In another embodiment, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising any of the nucleotide sequences as described above, or a complement thereof, and a pharmaceutically acceptable carrier.

In a related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of the inventive hEbola virus described above, and a pharmaceutically acceptable carrier. In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a protein extract of the inventive hEbola virus described above or a subunit thereof; and a pharmaceutically acceptable carrier. In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 or a complement thereof, and a pharmaceutically acceptable carrier. In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a nucleic acid molecule comprising the inventive nucleotide sequence as described above or a complement thereof, and a pharmaceutically acceptable carrier. In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of any of the inventive polypeptides described above.

In another aspect, the present invention provides pharmaceutical compositions comprising antiviral agents of the present invention and a pharmaceutically acceptable carrier. In a specific embodiment, the antiviral agent of the invention is an antibody that immunospecifically binds hEbola virus.
virus or any hEbola epitope. In another specific embodiment, the antiviral agent is a polypeptide or protein of the present invention or nucleic acid molecule of the invention.

[0029] In a related aspect, the invention provides a pharmaceutical composition comprising a prophylactically or therapeutically effective amount of an anti-hEbola EboBun agent and a pharmaceutically acceptable carrier.

[0030] The invention also provides kits containing compositions and formulations of the present invention. Thus, in another aspect, the invention provides a kit comprising a container containing the inventive immunogenic formulation described above. In another aspect, the invention provides a kit comprising a container containing the inventive vaccine formulation described above. In another aspect, the invention provides a kit comprising a container containing the inventive pharmaceutical composition described above. In another aspect, the invention provides a kit comprising a container containing the inventive vaccine formulation described above. In another invention provides a method for identifying a subject infected with the inventive hEbola virus described above, comprising: (a) obtaining total RNA from a biological sample obtained from the subject; (b) reverse transcribing the total RNA to obtain cDNA; and (c) amplifying the cDNA using a set of primers derived from a nucleotide sequence of the inventive hEbola virus described above.

[0031] The invention further relates to the use of the sequence information of the isolated virus for diagnostic and therapeutic methods.

[0032] In another aspect, the present invention provides methods for screening antiviral agents that inhibit the infectivity or replication of hEbola virus or variants thereof.

[0033] The invention further provides methods of preparing recombinant or chimeric forms of hEbola.

**BRIEF DESCRIPTION OF THE DRAWINGS**

[0034] FIG. 1 represents a Phylogenetic tree comparing full-length genomes of Ebola virus and Marburg virus by Bayesian analysis:

[0035] FIG. 2 represents an alignment of genomes of novel hEbola EboBun (SEQ ID NO: 1) referred to below as “EboBun Bundibugyo” or “EboBun”, and hEbola Zaire (SEQ ID NO: 20) referred to below as “Ebola Zaire ’76” or “EboZ” and hEbola Ivory Coast (SEQ ID NO: 10) also referred to below as “EboIC”.

**DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

[0036] It is to be understood that the present invention is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0037] Due to the sequence divergence of EboBun relative to all previously recognized ebolaviruses, the present invention has utility in design of diagnostic assays to monitor Ebola HF disease in humans and animals, and develop effective antivirals and vaccines.

[0038] The EboBun virus of the present invention is genetically distinct, differing by more than 30% at the genome level from all other known ebolavirus species. The unique nature of this virus created challenges for traditional filovirus molecular based diagnostic assays and genome sequencing approaches. Instead, over 70% of the virus genome was sequenced using a recently developed random-primed pyrosequencing approach which allowed the rapid development of molecular detection assay which were deployed in the disease outbreak response. This random-primed pyrosequencing draft sequence allowed faster completion of the whole genome sequence using traditional primer walking approach and confirmation that the EboBun virus represented a new ebolavirus species.

Definitions

[0039] The definitions herein provided are operative throughout the entire description of the invention set forth herein, including the Summary of the Invention.

[0040] The term “an antibody or an antibody fragment that immunospecifically binds a polypeptide of the invention” as used herein refers to an antibody or a fragment thereof that immunospecifically binds to the polypeptide encoded by the nucleotide sequence of SEQ ID NO: 1 (EboBun), or a fragment thereof, and does not non-specifically bind to other polypeptides. An antibody or a fragment thereof that immunospecifically binds to the polypeptide of the invention may cross-react with other antigens. Preferably, an antibody or a fragment thereof that immunospecifically binds to a polypeptide of the invention does not cross-react with other antigens. An antibody or a fragment thereof that immunospecifically binds to the polypeptide of the invention can be identified by, for example, immunoassays or other techniques known to those skilled in the art, or otherwise as described herein.

[0041] An “isolated” or “purified” polypeptide or protein is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language “substantially free of cellular material” includes preparations of a polypeptide/protein in which the polypeptide/protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, a polypeptide/protein that is substantially free of cellular material includes preparations of the polypeptide/protein having less than about 30%, 20%, 10%, 5%, 2.5%, or 1% (by dry weight) of contaminating protein. When the polypeptide/protein is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation.

[0042] When polypeptide/protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, i.e., it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly, such preparations of the polypeptide/protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than polypeptide/protein fragment of interest. In a preferred embodiment of the present invention, polypeptides/proteins are isolated or purified.

[0043] An “isolated” nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Moreover, an “isolated” nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In a preferred embodiment of the invention, nucleic acid molecules encod-
polypeptides/proteins of the invention are isolated or purified. The term "isolated" nucleic acid molecule does not include a nucleic acid that is a member of a library that has not been purified away from other library clones containing other nucleic acid molecules.

The term "portion" or "fragment" as used herein includes the specified fragment lengths, and all integers in between, inclusive of the specified end points in a specified range, and inclusive of any length up to the full length of a protein, polypeptide, or nucleic acid.

The term "having a biological activity of the protein" or "having biological activities of the polypeptides of the invention" refers to the characteristics of the polypeptides or proteins having a common biological activity, similar or identical structural domain, and/or having sufficient amino acid identity to the polypeptide encoded by the nucleotide sequence of SEQ ID NO: 1 (EboBun). Such common biological activities of the polypeptides of the invention include antigenticity and immunogenicity.

The term "under stringent condition" refers to hybridization and washing conditions under which nucleotide sequences having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% identity to each other remain hybridized to each other. Such hybridization conditions are described in, for example but not limited to, Current Protocols in Molecular Biology, John Wiley & Sons, NY (1989), 6.3.1-6.3.6; Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., NY (1986), pp. 75-78, and 84-87; and Molecular Cloning, Cold Spring Harbor Laboratory, NY (1982), pp. 387-389, and are well known to those skilled in the art. A preferred, non-limiting example of stringent hybridization conditions is hybridization in 6x sodium chloride/sodium citrate (SSC), 0.5% SDS at about 68°C followed by one or more washes in 2x SSC, 0.5% SDS at room temperature. Another preferred, non-limiting example of stringent hybridization conditions is hybridization in 6x SSC at about 45°C followed by one or more washes in 0.2x SSC, 0.1% SDS at about 50-65°C.

The term "variant" as used herein refers either to a naturally occurring genetic mutant of hB:olga EboBun, or hB:olga EboIC or a recombinantly prepared variant of these H:B:ola species, each of which contain one or more mutations in its genome compared to the H:B:ola of SEQ ID NO: 1 or 10. The term "variant" may also refer either to a naturally occurring variation of a given polypeptide or a recombinantly prepared variation of a given polypeptide or protein in which one or more amino acid residues have been modified by amino acid substitution, addition, or deletion.

"Homology" refers to sequence similarity or, alternatively, sequence identity, between two or more nucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to nucleotide sequences, refer to the percentage of identical nucleotide matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using one or more computer algorithms or programs known in the art or described herein. For example, percent identity can be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASEGNE software package, a suite of molecular biological analysis programs (DNASTAR, Madison, Wis.). CLUSTAL V is described in Higgins, D.G. and P. M. Sharp (1989; CABIOS 5:151-153) and in Higgins, D.G. et al. (1992; CABIOS 8:189-191). For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms which can be used is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, Md., and on the NCBI World Wide Web site available on the Internet. The BLAST software suite includes various sequence analysis programs including "b1asta," that is used to align a known nucleotide sequence with other nucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively on the Internet via the NCBI World Wide Web site as well. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (described below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (Apr. 21, 2000) set at default parameters. Such default parameters may be, for example: Matrix:BLOSUM62; Reward for match: 1; Penalty for mismatch: -2; Open Gap: 5 and Extension Gap: 2 penalties; Gapdrop-off: 50; Expect: 10; Word Size: 11; Filter: on.

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or sequence listing, may be used to describe a length over which percentage identity may be measured.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of identical residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide. The phrases "percent similarity" and "% similarity," as applied to polypeptide sequences, refer to the percentage of residue matches, including identical residue matches and conservative substitutions, between at least two polypeptide sequences aligned using a standardized algorithm. In contrast, conservative substitutions are not included in the calculation of percent identity between polypeptide sequences.
Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGA-LIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktule=1, gap penalty=–3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default residue weight table.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version 2.0.12 (Apr. 21, 2000) with blastsp set at default parameters. Such default parameters may be, for example: Matrix: BLOSUM62; Open Gap: 11 and Extension Gap: 1 penalties; GapXdrop-off: 50; Expect: 10; Word Size: 3; Filter: on.

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or sequence listing, may be used to describe a length over which percentage identity may be measured.

The term “agent” encompasses any chemical, biochemical, or biological molecule; such as small molecules, proteins, polypeptides, antibodies, nucleic acid molecules including DNA or RNA, and the like.

Methods and Compositions Related to the Inventive hEbola

The present invention is based upon the isolation and identification of a new human Ebola virus species, EboBun and the sequencing of the only other known West African Ebola species EboIC. EboBun was isolated from the patients suffering from hemorrhagic fever in a recent outbreak in Uganda. The isolated virus is a member of the Filoviridae family, a family of negative sense RNA viruses. Accordingly, the invention relates to the isolated EboBun or EboIC virus that morphologically and phylogenetically relates to known members filoviridae.

In another aspect, the invention provides an isolated hEbola virus including a nucleic acid molecule with a nucleotide sequence that is preferably: a) a nucleotide sequence set forth in SEQ ID NO: 1; b) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 1 under stringent conditions; or c) a nucleotide sequence that has at least 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity to the SEQ ID NO: 1. In one embodiment of the present invention, the hEbola virus is killed. In another, the virus is attenuated. In another, the infectivity of the attenuated hEbola virus is reduced. In another, the infectivity is reduced by at least 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 250-fold, 500-fold, or 10,000-fold. In another, the replication ability of the attenuated hEbola virus is reduced. In another, the replication ability of the attenuated hEbola virus is reduced by at least 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 250-fold, 500-fold, 1,000-fold, or 10,000-fold. In another, the assembling ability of the attenuated hEbola virus is reduced. In another, the assembling ability of the attenuated virus is reduced by at least 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 250-fold, 500-fold, 1,000-fold, or 10,000-fold.
sequence selected from the group consisting of: a) an amino acid sequence set forth in SEQ ID NO: 2, 3, 4, 5, 6, 7, 8, or 9; 11, 12, 13, 14, 15, 16, 17, 18 or 19; and b) an amino acid sequence that has 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% homology to the amino acid sequence according to a). In another, the isolated polypeptide comprises the amino acid sequence having at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 50, 60, 70, 80, 90, 100, 150, 200, 210, 220, 230, 240 or 250 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 5 or 18 (VP34); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 50, 60, 70, 80, 90, 100, 150, 200, 210, 220, 230, 240, 250, 260, 270, 280 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 6 or 17 (VP30); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 310, or 320 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 8 or 13 (VP40); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 310, 320, 330, 340, 350, 360, or 370 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 7 or 12 (VP35); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 310, 320, 330, 340, 350, 360, or 370 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 4 or 15 (SGP); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 610, 620, 630, 640, 650, 660, or 670 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 9 or 14 (GP); 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2150, 2160, 2170, 2180, 2190, or 2200 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 2 or 19 (I).

In another aspect, the invention relates to the use of an isolated West African hEbola virus for diagnostic and therapeutic methods. In one embodiment, the invention provides a method of detecting in a biological sample an antibody immunospecific for the hEbola virus using the inventive isolated hEbola virus described herein, or any of the inventive proteins or polypeptides as described herein. In another specific embodiment, the invention provides a method of screening for an antibody which immunospecifically binds and neutralizes hEbola EboBun or EboIC or a combination thereof. Such an antibody is useful for a passive immunization or immunotherapy of a subject infected with hEbola.

In another aspect, the invention provides an isolated antibody or an antigen-binding fragment thereof which immunospecifically binds to a West African genus hEbola virus of the invention described above, and illustratively including EboBun or EboIC. In one embodiment of the present invention, the isolated antibody or an antigen-binding fragment thereof neutralizes a West African genus hEbola virus. In another, the isolated antibody or an antigen-binding fragment thereof immunospecifically binds to the inventive polypeptide described above. The invention further provides antibodies that specifically bind a polypeptide of the invention encoded by the nucleotide sequence of SEQ ID NOs: 1 (EboBun) or 10 (EboIC), a fragment thereof, or encoded by a nucleic acid comprising a nucleotide sequence that hybridizes under stringent conditions to the nucleotide sequence of SEQ ID NOs: 1 (EboBun) or 10 (EboIC) and/or any hEbola EboBun epitope, having one or more biological activities of a polypeptide of the invention. These polypeptides include those shown in SEQ ID NOs: 2-9, 59, and 11-19. Such antibodies include, but are not limited to, polyclonal, monoclonal, bi-specific, multi-specific, human, humanized, chimeric antibodies, single chain antibodies, Fab fragments, F(ab')2 fragments, disulfide-linked Fvs, intrabodies and fragments containing either a VL or VH domain or even a complementary determining region (CDR) that specifically binds to a polypeptide of the invention.

In another aspect, the invention provides methods for detecting the presence, activity or expression of the hEbola virus of the invention in a biological material, such as cells, blood, saliva, urine, and so forth. The increased or decreased activity or expression of the hEbola virus in a sample relative to a control sample can be determined by contacting the biological material with an agent which can detect directly or indirectly the presence, activity or expression of the hEbola virus. In one embodiment of the present invention, the detecting agents are the antibodies or nucleic acid molecules of the present invention. Antibodies of the invention can also be used to treat hemorrhagic fever.

In a related aspect, the invention provides a method for detecting the presence of the inventive hEbola virus described above in a biological sample, the method comprising: (a) contacting the sample with an agent that selectively binds to the hEbola virus; and (b) detecting whether the compound binds to the hEbola virus in the sample. In one embodiment of the present invention, the biological sample is selected from the group consisting of cells; blood; serum; plasma; feces; rectal, vaginal and conjunctival swabs. In another, the agent that binds to the virus is an antibody. In another, the agent that binds to the virus is a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 or a complement thereof. In another, the agent that binds to the virus is a nucleic acid molecule comprising a nucleotide sequence having at least 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2150, 2160, 2170, 2180, 2190, or 2200 contiguous amino acid residues of the amino acid sequence of SEQ ID NOs: 1 or 19 (I).

In another aspect, the invention provides a method for detecting the presence of the inventive polypeptide described above, in a biological sample, the method comprising: (a) contacting the biological sample with an agent that selectively binds to the polypeptide; and (b) detecting whether the agent binds to the polypeptide in the sample. In one embodiment of the present invention, the biological sample is selected from the group consisting of cells; blood; serum; plasma; feces; rectal, vaginal and conjunctival swabs. In another, the agent that binds to the polypeptide is an antibody or an antigen-binding fragment thereof.

In another aspect, the invention provides a method for detecting the presence of a first nucleic acid molecule derived from the inventive hEbola virus described above in a biological sample, the method includes (a) contacting the biological sample with an agent that selectively binds to the
nucleic acid; and (b) detecting whether the agent binds to the nucleotide in the sample. In one embodiment of the present invention, the agent that binds to the first nucleic acid molecule is a second nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1 or a complement thereof. In another embodiment, the second nucleic acid molecule comprises at least 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 4600, 4700, 4800, 4900, 5000, 5500, 5600, 5700, 5800, 5900, 6000, 6100, 6200, 6300, 6400, 6500, or 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NOs: 1 or 10, or a complement thereof.

[0070] In another aspect, the invention provides a method for propagating the hEbola virus in host cells comprising infecting the host cells with an inventive isolated West African hEbola virus described above, culturing the host cells to allow the virus to multiply, and harvesting the resulting virions. Also provided by the present invention are host cells infected with the inventive hEbola virus described above. In one embodiment of the present invention, the host cell is a primate cell.

[0071] In another aspect, the invention provides a method of detecting in a biological sample the presence of an antibody that immunospecifically binds hEbola virus, the method includes: (a) contacting the biological sample with the inventive host cell described above; and (b) detecting the antibody bound to the cell.

[0072] In another aspect, the invention provides vaccine preparations, including the inventive hEbola virus, including recombinant and chimeric forms of the virus, nucleic acid molecules comprised by the virus, or protein subunits of the virus. In one embodiment, the vaccine preparations of the present invention includes live but attenuated hEbola virus with or without pharmaceutically acceptable carriers, including adjuvants. In another, the vaccine preparations of the invention comprise an inactivated or killed hEbola EboRun virus, EboC virus, or a combination thereof, with or without pharmaceutically acceptable carriers, including adjuvants. Such attenuated or inactivated viruses may be prepared by a series of passages of the virus through the host cells or by preparing recombinant or chimeric forms of the virus. Accordingly, the present invention further provides methods of preparing recombinant or chimeric forms of the inventive hEbola viruses described herein.

[0073] In another specific embodiment, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of the inventive hEbola virus described above, and a pharmaceutically acceptable carrier. In another, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a protein extract of the inventive hEbola virus described above, or a subunit thereof; and a pharmaceutically acceptable carrier. In another aspect, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1 or 10, or a complement thereof, and a pharmaceutically acceptable carrier. In another, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising any of inventive the nucleotide sequences as described above, or a complement thereof, and a pharmaceutically acceptable carrier. In another aspect, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a protein extract of the inventive hEbola virus described above, or a subunit thereof; and a pharmaceutically acceptable carrier. In another aspect, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1 or 10, or a complement thereof, and a pharmaceutically acceptable carrier. In another aspect, the invention provides a vaccine formulation comprising a therapeutically or prophylactically effective amount of a nucleic acid molecule comprising any of inventive the nucleotide sequences as described above, or a complement thereof, and a pharmaceutically acceptable carrier.

[0074] In yet another specific embodiment, the vaccine preparations of the present invention comprise a nucleic acid or fragment of the hEbola virus, e.g., the virus having Accession No. 200706291, or nucleic acid molecules having the sequence of SEQ ID NOs: 1 or 10, or a fragment thereof. In another, the vaccine preparations comprise a polypeptide of the invention encoded by the nucleotide sequence of SEQ ID NOs: 1 or 10 or a fragment thereof. In a specific embodiment, the vaccine preparations comprise polypeptides of the invention as shown in SEQ ID NOs: 2-9, 59, or 11-19, or encoded by the nucleotide sequence of SEQ ID NOs: 1 or 10, or a fragment thereof.

[0075] Furthermore, the present invention provides methods for treating, ameliorating, managing or preventing hemorrhagic fever by administering the vaccine preparations or antibodies of the present invention alone or in combination with adjuvants, or other pharmaceutically acceptable excipients. Furthermore, the present invention provides methods for treating, ameliorating, managing, or preventing hemorrhagic fever by administering the inventive compositions and formulations including the vaccine preparations or antibodies of the present invention alone or in combination with antivirals [e.g., amantadine, rimantadine, gancyclovir, acyclovir, ribavirin, penciclovir, oseltamivir, foscamet, idoxuridine (AZT), didanosine (ddI), lamivudine (3TC), zalcitabine (ddC), stavudine (d4T), nevirapine, delavirdine, indinavir, ritonavir, vidarabine, nelfinavir, saquinavir, relenza, tamiflu, pleconaril, interferons, etc.], steroids and corticosteroids such as prednisone, cortisone, fluticasone and glucocorticoid, antibiotics, analgesics, bronchodilators, or other treatments for respiratory and/or viral infections.

[0076] In a related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of the inventive hEbola virus described above, and a pharmaceutically acceptable carrier.

[0077] In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a protein extract of the inventive hEbola virus described above or a subunit thereof, and a pharmaceutically acceptable carrier.

[0078] In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1, 10, a combination thereof, or a complement thereof, and a pharmaceutically acceptable carrier.

[0079] In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of a nucleic acid molecule comprising the
inventive nucleotide sequence as described above or a complement thereof, and a pharmaceutically acceptable carrier.

[0080] In another related aspect, the invention provides an immunogenic formulation comprising an immunogenically effective amount of any of the inventive polypeptides described above.

[0081] In another aspect, the present invention provides pharmaceutical compositions comprising antiviral agents of the present invention and a pharmaceutically acceptable carrier. In a specific embodiment, the antiviral agent of the invention is an antibody that immunospecifically binds hEbola virus or any hEbola epitope. In another specific embodiment, the antiviral agent is a polypeptide or protein of the present invention or nucleic acid molecule of the invention.

[0082] In a related aspect, the invention provides a pharmaceutical composition comprising a prophylactically or therapeutically effective amount of an anti-hEbola EboBun agent and a pharmaceutically acceptable carrier. In one embodiment of the present invention, the anti-hEbola EboBun agent is an antibody or an antigen-binding fragment thereof which immunospecifically binds to the hEbola virus of Deposit Accession No. 200706291, or polypeptides or protein derived therefrom. In another, the anti-hEbola agent is a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1, 10, a combination thereof, or a fragment thereof. In another, the anti-hEbola agent is a polypeptide encoded by a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOs: 1, 10, a combination thereof, or a fragment thereof having a biological activity of the polypeptide.

[0083] The invention also provides kits containing compositions and formulations of the present invention. Thus, in another aspect, the invention provides a kit comprising a container containing the inventive immunogenic formulation described above.

[0084] In another aspect, the invention provides a kit including a container containing the inventive vaccine formulation described above.

[0085] In another aspect, the invention provides a kit including a container containing the inventive pharmaceutical composition described above.

[0086] In another aspect, the invention provides a kit including a container containing the inventive vaccine formulation described above.

[0087] In another aspect, the invention provides a method for identifying a subject infected with the inventive hEbola virus described above, including: (a) obtaining total RNA from a biological sample obtained from the subject; (b) reverse transcribing the total RNA to obtain cDNA; and (c) amplifying the cDNA using a set of primers derived from a nucleotide sequence of the inventive hEbola virus described above.

[0088] In one embodiment of the present invention, the set of primers are derived from the nucleotide sequence of the genome of the hEbola virus of Deposit Accession No. 200706291. In another, the set of primers are derived from the nucleotide sequence of SEQ ID NOs: 1 or 10 or any of the inventive nucleotide sequences as described above, or a complement thereof.

[0089] The invention further relates to the use of the sequence information of the isolated virus for diagnostic and therapeutic methods. In a specific embodiment, the invention provides nucleic acid molecules which are suitable for use as primers consisting of or including the nucleotide sequence of SEQ ID NOs: 1 or 10, or a complement thereof, or at least a portion of the nucleotide sequence thereof. In another specific embodiment, the invention provides nucleic acid molecules which are suitable for hybridization to the inventive hEbola nucleic acid; including, but not limited to PCR primers, Reverse Transcriptase primers, probes for Southern analysis or other nucleic acid hybridization analysis for the detection of hEbola nucleic acids, e.g., consisting of or including the nucleotide sequence of SEQ ID NOs: 1, 10 a combination thereof, a complement thereof, or a fragment thereof. The invention further encompasses chimeric or recombinant viruses encoded in whole or in part by the nucleotide sequences.

[0090] In another aspect, the present invention provides methods for screening antiviral agents that inhibit the infectivity or replication of hEbola virus or variants thereof.

[0091] The invention further provides methods of preparing recombinant or chimeric forms of hEbola.

[0092] In another aspect, the invention provides vaccine preparations including the hEbola virus, including recombinant and chimeric forms of the virus, or subunits of the virus. The present invention encompasses recombinant or chimeric viruses encoded by viral vectors derived from the genome of the inventive hEbola virus described herein or natural variants thereof. In a specific embodiment, a recombinant virus is one derived from the hEbola virus of Deposit Accession No. 200706291. It is recognized that natural variants of the inventive hEbola viruses described herein comprise one or more mutations, including, but not limited to, point mutations, rearrangements, insertions, deletions etc., in the genomic sequence. It is recognized that the mutations may or may not result in a phenotypic change.

[0093] In another specific embodiment, a chimeric virus of the invention is a recombinant hEbola EboBun or EboIC virus which further comprises a heterologous nucleotide sequence. In accordance with the invention, a chimeric virus may be encoded by a nucleotide sequence in which heterologous nucleotide sequences have been added to the genome or in which endogenous or native nucleotide sequences have been replaced with heterologous nucleotide sequences.

[0094] According to the present invention, the chimeric viruses are encoded by the viral vectors of the invention which further comprise a heterologous nucleotide sequence. In accordance with the present invention a chimeric virus is encoded by a viral vector that may or may not include nucleic acids that are non-native to the viral genome. In accordance with the invention a chimeric virus is encoded by a viral vector to which heterologous nucleotide sequences have been added, inserted or substituted for native or non-native sequences. In accordance with the present invention, the chimeric virus may be encoded by nucleotide sequences derived from different species or variants of hEbola virus. In particular, the chimeric virus is encoded by nucleotide sequences that encode antigenic polypeptides derived from different species or variants of hEbola virus.

[0095] A chimeric virus may be of particular use for the generation of recombinant vaccines protecting against two or more viruses (Tao et al., J. Virol. 72, 2955-2961; Durbin et al., 2000, J. Virol. 74, 6821-6831; Skiadopoulos et al., 1998, J. Virol. 72, 1762-1768 (1998); Teng et al., 2000, J. Virol. 74, 9317-9321). For example, it can be envisaged that a virus vector derived from the hEbola virus expressing one or more proteins of variants of hEbola virus including hEbola EboBun, or vice versa, will protect a subject vaccinated with
such vector against infections by both the native hEbola and the variant. Attenuated and replication-defective viruses may be of use for vaccination purposes with live vaccines as has been suggested for other viruses. (See, for example, PCT WO 02/057302, at pp. 6 and 23; and United States Patent Application Publication 2008/0069838 incorporated by reference herein).

[0096] In accordance with the present invention the heterologous sequence to be incorporated into the viral vectors encoding the recombinant or chimeric viruses of the invention include sequences obtained or derived from different species or variants of hEbola.

[0097] In certain embodiments, the chimeric or recombinant viruses of the invention are encoded by viral vectors derived from viral genomes wherein one or more sequences, intergenic regions, termini sequences, or portions or entire ORF have been substituted with a heterologous or non-native sequence. In certain embodiments of the invention, the chimeric viruses of the invention are encoded by viral vectors derived from viral genomes wherein one or more heterologous sequences have been inserted or added to the vector.

[0098] The selection of the viral vector may depend on the species of the subject that is to be treated or protected from a viral infection. If the subject is human, then an attenuated hEbola virus can be used to provide the antigenic sequences.

[0099] In accordance with the present invention, the viral vectors can be engineered to provide antigenic sequences which confer protection against infection by the inventive hEbola and natural variants thereof. The viral vectors may be engineered to provide one, two, three or more antigenic sequences. In accordance with the present invention the antigenic sequences may be derived from the same virus, from different species or variants of the same type of virus, or from different viruses.

[0100] The expression products and/or recombinant or chimeric virions obtained in accordance with the invention may advantageously be utilized in vaccine formulations. The expression products and chimeric virions of the present invention may be engineered to create vaccines against a broad range of pathogens, including viral and bacterial antigens, tumor antigens, allergen antigens, and auto antigens involved in autoimmune disorders. One way to achieve this goal involves modifying existing hEbola genes to contain foreign sequences in their respective external domains. Where the heterologous sequences are epitopes or antigens of pathogens, these chimeric viruses may be used to induce a protective immune response against the disease agent from which these determinants are derived. In particular, the chimeric virions of the present invention may be engineered to create vaccines for the protection of a subject from infections with hEbola virus and variants thereof.

[0101] Thus, the present invention further relates to the use of viral vectors and recombinant or chimeric viruses to formulate vaccines against a broad range of viruses and/or antigens. The present invention also encompasses recombinant viruses including a viral vector derived from the hEbola or variants thereof which contains sequences which result in a virus having a phenotype more suitable for use in vaccine formulations, e.g., attenuated phenotype or enhanced antigenicity. The mutations and modifications can be in coding regions, in intergenic regions and in the leader and trailer sequences of the virus.

[0102] The invention provides a host cell including a nucleic acid or a vector according to the invention. Plasmid or viral vectors containing the polymerase components of hEbola virus are generated in prokaryotic cells for the expression of the components in relevant cell types (bacteria, insect cells, eukaryotic cells). Plasmids or viral vectors containing full-length or partial copies of the hEbola genome will be generated in prokaryotic cells for the expression of viral nucleic acids in vitro or in vivo. The latter vectors optionally contain other viral sequences for the generation of chimeric viruses or chimeric virus proteins, optionally lack parts of the viral genome for the generation of replication-defective virus, and optionally contain mutations, deletions or insertions for the generation of attenuated viruses. In addition, the present invention provides a host cell infected with hEbola virus of Deposit Accession No. 200706291.

[0103] Infectious copies of West African hEbola (being wild type, attenuated, replication-defective or chimeric) are optionally produced upon co-expression of the polymerase components according to the state-of-the-art technologies described above.

[0104] In addition, eukaryotic cells, transiently or stably expressing one or more full-length or partial hEbola proteins are optionally used. Such cells are preferably made by transfection (proteins or nucleic acid vectors), infection (viral vectors) or transduction (viral vectors) and are useful for complementation of mentioned wild type, attenuated, replication-defective or chimeric viruses.

[0105] The viral vectors and chimeric viruses of the present invention optionally modulate a subject's immune system by stimulating a humoral immune response, a cellular immune response or by stimulating tolerance to an antigen. As used herein, a subject means: humans, primates, horses, cows, sheep, pigs, goats, dogs, cats, avian species and rodents.

Formulation of Vaccines and Antivirals

[0106] In a preferred embodiment, the invention provides a proteinaceous molecule or hEbola virus specific viral protein or functional fragment thereof encoded by a nucleic acid according to the invention. Useful proteinaceous molecules are for example derived from any of the genes or genomic fragments derivable from the virus according to the invention, preferably the GP, L, NP, sGP, VP24, VP30, VP35, and VP 40 proteins described herein. Such molecules, or antigenic fragments thereof, as provided herein, are for example useful in diagnostic methods or kits and in pharmaceutical compositions such as subunit vaccines. Particularly useful are polypeptides encoded by the nucleotide sequence of SEQ ID NOs: 1 or 10; or antigenic fragments thereof for inclusion as antigen or subunit immunogen, but inactivated whole virus can also be used. Particularly useful are also those proteinaceous substances that are encoded by recombinant nucleic acid fragments of the hEbola genome, of course preferred are those that are within the preferred bounds and metes of ORFs, in particular, for eliciting hEbola specific antibody or T cell responses, whether in vivo (e.g. for protective or therapeutic purposes or for providing diagnostic antibodies) or in vitro (e.g. by phage display technology or another technique useful for generating synthetic antibodies).

[0107] It is recognized that numerous variants, analogues, or homologues of Ebo36nu polypeptides are within the scope of the present invention including amino acid substitutions, alterations, modifications, or other amino acid changes that increase, decrease, or do not alter the function or immunogenic propensity of the inventive immunogen or vaccine. Several post-translational modifications are similarly envi-
sioned as within the scope of the present invention illustra-
tively including incorporation of a non-naturally occurring
amino acid(s), phosphorylation, glycosylation, sulfation, and
addition of pendant groups such as biotinylation, fluoro-
phores, lumiphores, radioactive groups, antigens, or other
molecules.

Methods of expressing and purifying natural or re-
combinant peptides and proteins are well known in the art.
Illustratively, peptides and proteins are recombinantly
expressed in eukaryotic cells. Exemplary eukaryotic cells
include yeast, HeLa cells, 293 cells, COS cells, Chinese ham-
eroey cells (CHO), and many other cell types known in
the art. Both eukaryotic and prokaryotic expression systems
and cells are available illustratively from Invitrogen Corp.,
Carlsbad, Calif. It is appreciated that cell-free expression
systems are similarly operable.

In a preferred embodiment an immunogenic polypeptide
is a full length EboBun protein. Preferably, an
immunogen is a full length EboBun protein of SEQ ID NOs:
2-9 or 59, or EboC SEQ ID NOs: 11-19, or a fragment
thereof as described herein. Preferably, an immunogen is
has a minimum of 5 amino acids. As used herein an immunogen
is preferably a polypeptide. In the context of an immunogenic
polypeptide the terms immunogen, polypeptide, and antigen
are used interchangeably.

Modifications and changes can be made in the struc-
ture of the inventive immunogens that are the subject of the
application and still obtain a molecule having similar or
improved characteristics as the wild-type sequence (e.g., a
conservative amino acid substitution). For example, certain
amino acids are optionally substituted for other amino acids
in a sequence without appreciable loss of immunogenic activ-
ity. Because it is the interactive capacity and nature of a
polypeptide that defines that polypeptide’s biological func-
tional activity, certain amino acid sequence substitutions
can be made in a polypeptide sequence and nevertheless obtain
a polypeptide with like or improved properties. Optionally,
a polypeptide is used that has less or more immunogenic activ-
ity compared to the wild-type sequence.

In making such changes, the hydropathic index of
amino acids is preferably considered. The importance of the
hydropathic amino acid index in conferring interactive bio-
logic function on a polypeptide is generally understood in the
art. It is known that certain amino acids can be substituted for
other amino acids having a similar hydropathic index or score
and still result in a polypeptide with similar biological activ-
ity. Each amino acid has been assigned a hydropathic index
on the basis of its hydrophobicity and charge characteristics.
Those indices are: isoleucine (+4.5); valine (+4.2); leucine
(+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine
(-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); pro-
line (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and
arginine (-4.5).

It is believed that the relative hydropathic character
of the amino acid determines the secondary structure of the
resultant polypeptide, which in turn defines the interaction of
the polypeptide with other molecules, such as enzymes,
substrates, receptors, antibodies, antigens, and the like. It is
known in the art that an amino acid can be substituted by
another amino acid having a similar hydropathic index and
still obtain a functionally equivalent immunogen. In such
changes, the substitution of amino acids whose hydropathic
indices are within ±2 is preferred, those within ±1 are par-
icularly preferred, and those within ±0.5 are even more par-
icularly preferred.

As outlined above, amino acid substitutions are gen-
erally based on the relative similarity of the amino acid side-
chain substituents, for example, their hydrophobicity, hydro-
philicity, charge, size, and the like. Exemplary substitutions
that take various of the foregoing characteristics into consid-
eration are well known to those of skill in the art and include
(original residues: exemplar substitution): (Ala: Gly, Ser),
(Arg: Lys), (Asp: Glu, His), (Asp: Glu, Cys, Ser), (Gln: Asn),
(Glu: Asp), (Gly: Ala), (His: Asn, Gln), (Ile: Leu, Val), (Leu:
Ile, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser),
(Trp: Tyr), (Tyr: Trp, Phe), and (Val: Ile, Leu). Embodiments of
this disclosure thus contemplate functional or biological
equivalents of a polypeptide and immunogen as set forth
above. In particular, embodiments of the polypeptides and
immunogens optionally include variants having about 50%,
60%, 70%, 80%, 90%, and 95% sequence identity to the
polypeptide of interest.

The invention provides vaccine formulations for the
prevention and treatment of infections with hEBoLA virus. In
certain embodiments, the vaccine of the invention comprises
recombinant and chimeric viruses of the hEBoLA virus.
In certain embodiments, the virus is attenuated.

In another embodiment of this aspect of the inven-
tion, inactivated vaccine formulations are prepared using con-
ventional techniques to “kill” the chimeric viruses. Inacti-
vated vaccines are “dead” in the sense that their infectivity
has been destroyed. Ideally, the infectivity of the virus is
destroyed without affecting its immunogenicity. In order to
prepare inactivated vaccines, the chimeric virus may be
grown in cell culture or in the allantois of the chick embryo,
purified by zonal ultracentrifugation, inactivated by formal-
dehyde or β-propiolactone, and pooled. The resulting vaccine
is usually inoculated intramuscularly or intranasally.

Inactivated viruses are optionally formulated with a
suitable adjuvant in order to enhance the immunological
response. Such adjuvants illustratively include but are not
limited to mineral gels, e.g., aluminum hydroxide; surface
active substances such as lyssolecithin, pluronic polylols,
polyanions; peptides; oil emulsions; and potentially useful
human adjuvants such as BCG and Corynebacterium parvum.

In another aspect, the present invention also pro-
vides DNA vaccine formulations including a nucleic acid or
fragment of the inventive hEBoLA virus, e.g., the virus having
Accession No. 200706291, or nucleic acid molecules having
the sequence of SEQ ID NOS: 1 or 10, or a fragment thereof.
In another specific embodiment, the DNA vaccine formulat-
ions of the present invention comprise a nucleic acid or
fragment thereof encoding the antibodies which immunospe-
cifically bind hEBoLA viruses. In DNA vaccine formulations,
a vaccine DNA comprises a viral vector, such as that derived
from the hEBoLA virus, bacterial plasmid, or other expression
vector, bearing an insert including a nucleic acid molecule of
the present invention operably linked to one or more control
elements, thereby allowing expression of the vaccinating pro-
teins encoded by the nucleic acid molecule in a vaccinated
subject. Such vectors can be prepared by recombinant DNA
technology as recombinant or chimeric viral vectors carrying
a nucleic acid molecule of the present invention.

A nucleic acid as used herein refers to single- or
double-stranded molecules which are optionally DNA,
including the nucleotide bases A, T, C and G, or RNA, including the bases A, U (substitutes for T), C, and G. The nucleic acid may represent a coding strand or its complement. Nucleic acids are optionally identical in sequence to the sequence which is naturally occurring or include alternative codons which encode the same amino acid as that which is found in the naturally occurring sequence. Furthermore, nucleic acids optionally include codons which represent conservative substitutions of amino acids as are well known in the art.

[0119] As used herein, the term “isolated nucleic acid” means a nucleic acid separated or substantially free from at least some of the other components of the naturally occurring organism, for example, the cell structural components commonly found associated with nucleic acids in a cellular environment and/or other nucleic acids. The isolation of nucleic acids is illustratively accomplished by techniques such as cell lysis followed by phenol plus chloroform extraction, followed by ethanol precipitation of the nucleic acids. The nucleic acids of this invention are illustratively isolated from cells according to methods well known in the art for isolating nucleic acids. Alternatively, the nucleic acids of the present invention are optionally synthesized according to standard protocols well described in the literature for synthesizing nucleic acids. Modifications to the nucleic acids of the invention are also contemplated, provided that the essential structure and function of the peptide or polypeptide encoded by the nucleic acid are maintained.

[0120] The nucleic acid encoding the peptide or polypeptide of this invention is optionally part of a recombinant nucleic acid construct comprising any combination of restriction sites and/or functional elements as are well known in the art which facilitate molecular cloning and other recombinant DNA manipulations. Thus, the present invention further provides a recombinant nucleic acid construct including a nucleic acid encoding a polypeptide of this invention.

[0121] Generally, it may be more convenient to employ as the recombinant polynucleotide a cDNA version of the polynucleotide. It is believed that the use of a cDNA version will provide advantages in that the size of the gene will generally be much smaller and more readily employed to transfect the targeted cell than will a genome gene, which will typically be up to an order of magnitude larger than the cDNA gene. However, the inventor does not exclude the possibility of employing a genomic version of a particular gene where desired.

[0122] As used herein, the terms “engineered” and “recombinant” cells are synonymous with “host” cells and are intended to refer to a cell into which an exogenous DNA segment or gene, such as a cDNA or gene has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced exogenous DNA segment or gene. A host cell is optionally a naturally occurring cell that is transformed with an exogenous DNA segment or gene or a cell that is not modified. A host cell preferably does not possess a naturally occurring gene encoding RSV G protein. Engineered cells are, thus, cells having a gene or genes introduced through the hand of man. Recombinant cells illustratively include those having an introduced cDNA or genomic DNA, and also include genes positioned adjacent to a promoter not naturally associated with the particular introduced gene.

[0123] To express a recombinant encoded polypeptide in accordance with the present invention one optionally prepares an expression vector that comprises a polynucleotide under the control of one or more promoters. To bring a coding sequence “under the control of” a promoter, one positions the 5’ end of the translational initiation site of the reading frame generally between about 1 and 50 nucleotides “downstream” of (i.e., 3’ of) the chosen promoter. The “upstream” promoter stimulates transcription of the inserted DNA and promotes expression of the encoded recombinant protein. This is the meaning of “recombinant expression” in the context used here.

[0124] Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order to achieve protein or peptide expression in a variety of host-expression systems. Cell types available for expression include, but are not limited to, bacteria, such as E. coli and B. subtilis transformed with recombinant phage DNA, plasmid DNA or cosmid DNA expression vectors.

[0125] Certain examples of prokaryotic hosts illustratively include E. coli strain RR1, E. coli LE392, E. coli B, E. coli 1776 (ATCC No. 31537) as well as E. coli W3110 (F-, lambda-, prototrophic, ATCC No. 273325); bacilli such as Bacillus subtilis; and other enterobacteria such as Salmonella typhimurium, Serratia marcescens, and various Pseudomonas species.

[0126] In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, E. coli is often transformed using pBR322, a plasmid derived from an E. coli species. Plasmid pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage may also contain, or be modified to contain, promoters that can be used by the microbial organism for expression of its own proteins.

[0127] In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism are optionally used as transforming vectors in connection with these hosts. For example, the phage lambda is optionally utilized in making a recombinant phage vector that can be used to transform host cells, such as E. coli LE392.

[0128] Further useful vectors include pIN vectors and pGEX vectors, for use in generating glutathione S-transferase (GST) soluble fusion proteins for later purification and separation or cleavage. Other suitable fusion proteins are those with β-galactosidase, ubiquitin, or the like.

[0129] Promoters that are most commonly used in recombinant DNA construction include the β-lactamase (penicillinase), lactose and tryptophan (trp) promoter systems. While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling those of skill in the art to ligate them functionally with plasmid vectors.

[0130] For expression in Saccharomyces, the plasmid YRp7, for example, is commonly used. This plasmid contains the trp1 gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1. The presence of the trp1 lesion as a characteristic of the yeast host cell genome
then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

[0131] Suitable promoting sequences in yeast vectors illustratively include the promoters for 3-phosphoglycerate kinase or other glycolytic enzymes, such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucone isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also preferably ligated into the expression vector 3’ of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination.

[0132] Other suitable promoters, which have the additional advantage of transcription controlled by growth conditions, illustratively include the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphosphate dehydrogenase, and enzymes responsible for malolase and galactose utilization.

[0133] In addition to microorganisms, cultures of cells derived from multicellular organisms are also operable as hosts. In principle, any such cell culture is operable, whether from vertebrate or invertebrate culture. In addition to mammalian cells, these include insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus); and plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing one or more coding sequences.

[0134] In a useful insect system, Autographica californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in Spodoptera frugiperda cells. The isolated nucleic acid coding sequences are cloned into non-essential regions (for example the polyhedron gene) of the virus and placed under control of an AcNPV promoter (for example, the polyhedron promoter). Successful insertion of the coding sequences results in the inactivation of the polyhedron gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedron gene). These recombinant viruses are then used to infect Spodoptera frugiperda cells in which the inserted gene is expressed (e.g., U.S. Pat. No. 4,215,051).

[0135] Examples of useful mammalian host cell lines include VERO and Hela cells, Chinese hamster ovary (CHO) cell lines, W138, BHK, COS-7, 293, HepG2, NIH3T3, RIN and MDCK cell lines. In addition, a host cell is preferably chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the encoded protein.

[0136] Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems are preferably chosen to ensure the correct modification and processing of the foreign protein expressed. Expression vectors for use in mammalian cells ordinarily include an origin of replication (as necessary); a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences. The origin of replication is preferably provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

[0137] The promoters are optionally derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adeno virus late promoter, the vaccinia virus 7.5K promoter). Further, it is also possible, and may be desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

[0138] A number of viral based expression systems are operable herein, for example, commonly used promoters are derived from polyoma, Adenovirus 2, Adenovirus 5, cytomegalovirus and Simian Virus 40 (SV40). The early and late promoters of SV40 virus are useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments are also operable, particularly when there is included the approximately 250 bp sequence extending from the HindIII site to the BglII site located in the viral origin of replication.

[0139] In cases where an adenovirus is used as an expression vector, the coding sequences are preferably ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene is then optionally inserted into the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing proteins in infected hosts.

[0140] Specific initiation signals may also be required for efficient translation of the claimed isolated nucleic acid coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may additionally need to be provided. One of ordinary skill in the art would readily be capable of determining this need and providing the necessary signals. It is well known that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons are optionally of a variety of origins, both natural and synthetic. The efficiency of expression is optionally enhanced by the inclusion of appropriate transcription enhancer elements or transcription terminators.

[0141] In eukaryotic expression, one will also typically desire to incorporate into the transcriptional unit an appropriate polyadenylation site if one was not contained within the original cloned segment. Typically, the polyA addition site is placed about 30 to 2000 nucleotides “downstream” of the termination site of the protein at a position prior to transcription termination.

[0142] For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express constructs encoding proteins are engineered. Rather than using expression vectors that contain viral origins of replication, host cells are preferably transformed with vectors controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a
selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched medium, and then are switched to a selective medium. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci, which in turn can be cloned and expanded into cell lines.

A number of selection systems are illustratively used, including, but not limited to, the herpes simplex virus thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyltransferase genes, in tk-, hprt- or aprt- cells, respectively. Also, antimetabolite resistance is optionally used as the basis of selection for dhfr, which confers resistance to methotrexate; gpt, which confers resistance to mycophenolic acid; neo, which confers resistance to the aminoglycoside G-418; and hygro, which confers resistance to hygromycin. It is appreciated that numerous other selection systems are known in the art that are similarly operable in the present invention.

The nucleic acids encoding the peptides and polypeptides of this invention are optionally administered as nucleic acid vaccines. For the purposes of vaccine delivery, a nucleic acid encoding a peptide or polypeptide of this invention is preferably in an expression vector that includes viral nucleic acid including, but not limited to, vaccinia virus, adeno virus, retrovirus and/or aden-associated virus nucleic acid. The nucleic acid or vector of this invention is optionally in a liposome or a delivery vehicle which can be taken up by a cell via receptor-mediated or other type of endocytosis. The nucleic acid vaccines of this invention are preferably in a pharmaceutically acceptable carrier or administered with an adjuvant. The nucleic acids encoding the peptides and polypeptides of this invention can also be administered to cells in vivo or ex vivo.

It is contemplated that the isolated nucleic acids of the disclosure are optionally "overexpressed", i.e., expressed in increased levels relative to its natural expression in cells of its indigenous organism, or even relative to the expression of other proteins in the recombinant host cell. Such overexpression is assessed by a variety of methods illustratively including radio-labeling and/or protein purification. However, simple and direct methods are preferred, for example, those involving SDS/PAGE and protein staining or immunoblotting, followed by quantitative analyses, such as densitometric scanning of the resultant gel or blot. A specific increase in the level of the recombinant protein or peptide in comparison to the level in natural in transfected cells is indicative of overexpression, as is a relative abundance of the specific protein in relation to the other proteins produced by the host cell and, e.g., visible on a gel.


Many methods are optionally used to introduce the vaccine formulations described above. These include, but are not limited to, oral, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, and intranasal routes. Alternatively, in a preferred embodiment the chimeric virus vaccine formulation is introduced via the natural route of infection of the pathogen for which the vaccine is designed. The DNA vaccines of the present invention are optionally administered in saline solutions by injections into muscle or skin using a syringe and needle (Wolff, J. A. et al., 1990, Direct gene transfer into mouse muscle in vivo, Science 247:1465-1468; Raz, E., 1994, Intraderal gene immunization: The possible role of DNA uptake in the induction of cellular immunity to viruses, c. Natl. Acad. Sci. USA 91:9519-9523). Another way to administer DNA vaccines operable herein is called the "gene gun" method, whereby microscopic gold beads coated with the DNA molecules of interest are fired into the cells (Tang, D. et al., 1992, Genetic immunization is a simple method for eliciting an immune response, Nature 356:152-154). For general reviews of the methods for DNA vaccines, see Robinson, H. L., 1999, DNA vaccines: basic mechanism and immune responses (Review), Int. J. Mol. Med. 4(5):549-555; Barber, B., 1997, Introduction: Emerging vaccine strategies, Seminars in Immunology 9(5):269-270; and Robinson, H. L. et al., 1997, DNA vaccines, Seminars in Immunology 9(5):271-283.

Attention of hEbola Virus or Variants Thereof

The Ebola virus or variants thereof of the invention are optionally genetically engineered to exhibit an attenuated phenotype. In particular, the viruses of the invention exhibit an attenuated phenotype in a subject in which the virus is administered as a vaccine. Attenuation can be achieved by any method known to a skilled artisan. Without being bound by theory, the attenuated phenotype of the viruses of the invention is caused, e.g., by using a virus that naturally does not replicate well in an intended host species, for example, by reduced replication of the viral genome, by reduced ability of the virus to infect a host cell, or by reduced ability of the viral
proteins to assemble to an infectious viral particle relative to the wild type species of the virus.

[0149] The attenuated phenotypes of hEbola virus or variants thereof are optionally tested by any method known to the artisan. A candidate virus, for example, is optionally tested for its ability to infect a host or for the rate of replication in a cell culture system. In certain embodiments, growth curves at different temperatures are used to test the attenuated phenotype of the virus. For example, an attenuated virus is able to grow at 35°C, but not at 39°C or 40°C. In certain embodiments, different cell lines are used to evaluate the attenuated phenotype of the virus. For example, an attenuated virus may only be able to grow in monkey cell lines but not the human cell lines, or the achievable virus titers in different cell lines are different for the attenuated virus. In certain embodiments, viral replication in the respiratory tract of a small animal model, including but not limited to, hamsters, cotton rats, mice and guinea pigs, is used to evaluate the attenuated phenotypes of the virus. In other embodiments, the immune response induced by the virus, including but not limited to, the antibody titers (e.g., assayed by plaque reduction neutralization assay or ELISA) is used to evaluate the attenuated phenotypes of the virus. In a specific embodiment, the plaque reduction neutralization assay or ELISA is carried out at a low dose. In certain embodiments, the ability of the hEbola virus to elicit pathological symptoms in an animal model is tested. A reduced ability of the virus to elicit pathological symptoms in an animal model system is indicative of its attenuated phenotype. In a specific embodiment, the candidate viruses are tested in a monkey model for nasal infection, indicated by mucous production.

[0150] The viruses of the invention are optionally attenuated such that one or more of the functional characteristics of the virus are impaired. In certain embodiments, attenuation is measured in comparison to the wild type species of the virus from which the attenuated virus is derived. In other embodiments, attenuation is determined by comparing the growth of an attenuated virus in different host systems. Thus, for a non-limiting example, hEbola virus or a variant thereof is attenuated when grown in a human host if the growth of the hEbola or variant thereof in the human host is reduced compared to the non-attenuated hEbola or variant thereof.

[0151] In certain embodiments, the attenuated virus of the invention is capable of infecting a host, is capable of replicating in a host such that infectious viral particles are produced. In comparison to the wild type species, however, the attenuated species grows to lower titers or grows more slowly. Any technique known to the skilled artisan can be used to determine the growth curve of the attenuated virus and compare it to the growth curve of the wild type virus.

[0152] In certain embodiments, the attenuated virus of the invention (e.g., a recombinant or chimeric hEbola) cannot replicate in human cells as well as the wild type virus (e.g., wild type hEbola) does. However, the attenuated virus can replicate well in a cell line that lacks interferon functions, such as Vero cells.

[0153] In other embodiments, the attenuated virus of the invention is capable of infecting a host, of replicating in the host, and of causing proteins of the virus of the invention to be inserted into the cytoplasmic membrane, but the attenuated virus does not cause the host to produce new infectious viral particles. In certain embodiments, the attenuated virus infects the host, replicates in the host, and causes viral proteins to be inserted in the cytoplasmic membrane of the host with the same efficiency as the wild type hEbola. In other embodiments, the ability of the attenuated virus to cause viral proteins to be inserted into the cytoplasmic membrane into the host cell is reduced compared to the wild type virus. In certain embodiments, the ability of the attenuated hEbola virus to replicate in the host is reduced compared to the wild type virus. Any technique known to the skilled artisan can be used to determine whether a virus is capable of infecting a mammalian cell, of replicating within the host, and of causing viral proteins to be inserted into the cytoplasmic membrane of the host.

[0154] In certain embodiments, the attenuated virus of the invention is capable of infecting a host. In contrast to the wild type hEbola, however, the attenuated hEbola cannot be replicated in the host. In a specific embodiment, the attenuated hEbola virus can infect a host and can cause the host to insert viral proteins in its cytoplasmic membranes, but the attenuated virus is incapable of being replicated in the host. Any method known to the skilled artisan can be used to test whether the attenuated hEbola has infected the host and has caused the host to insert viral proteins in its cytoplasmic membranes.

[0155] In certain embodiments, the ability of the attenuated virus to infect a host is reduced compared to the ability of the wild type virus to infect the same host. Any technique known to the skilled artisan can be used to determine whether a virus is capable of infecting a host.

[0156] In certain embodiments, mutations (e.g., missense mutations) are introduced into the genome of the virus, for example, into the sequence of SEQ ID NOs: 1 or 10, or to generate a virus with an attenuated phenotype. Mutations (e.g., missense mutations) can be introduced into the structural genes and/or regulatory genes of the hEbola. Mutations are optionally additions, substitutions, deletions, or combinations thereof. Such variant of hEbola can be screened for a predicted functionality, such as infectivity, replication ability, protein synthesis ability, assembling ability, as well as cytopathic effect in cell cultures. In a specific embodiment, the missense mutation is a cold-sensitive mutation. In another embodiment, the missense mutation is a heat-sensitive mutation. In another embodiment, the missense mutation prevents a normal processing or cleavage of the viral proteins.

[0157] In other embodiments, deletions are introduced into the genome of the hEbola virus, which result in the attenuation of the virus.

[0158] In certain embodiments, attenuation of the virus is achieved by replacing a gene of the wild type virus with a gene of a virus of a different species, of a different subgroup, or of a different variant. In another aspect, attenuation of the virus is achieved by replacing one or more specific domains of a protein of the wild type virus with domains derived from the corresponding protein of a virus of a different species. In certain other embodiments, attenuation of the virus is achieved by deleting one or more specific domains of a protein of the wild type virus.

[0159] When a live attenuated vaccine is used, its safety should also be considered. The vaccine preferably does not cause disease. Any techniques known in the art for improving vaccine safety are operable in the present invention. In addition to attenuation techniques, other techniques are optionally used. One non-limiting example is to use a soluble heterologous gene that cannot be incorporated into the virion membrane. For example, a single copy of the soluble version
of a viral transmembrane protein lacking the transmembrane and cytosolic domains thereof is used.

Various assays are optionally used to test the safety of a vaccine. For example, sucrose gradients and neutralization assays are used to test the safety. A sucrose gradient assay is optionally used to determine whether a heterologous protein is inserted in a virion. If the heterologous protein is inserted in the virion, the virion is preferably tested for its ability to cause symptoms in an appropriate animal model since the virus may have acquired new, possibly pathological, properties.

5.4 Adjuvants and Carrier Molecules

hEbola-associated antigens are administered with one or more adjuvants. In one embodiment, the hEbola-associated antigen is administered together with a mineral salt adjuvant or mineral salt gel adjuvant. Such mineral salt and mineral salt gel adjuvants include, but are not limited to, aluminum hydroxide (ALHYDROGEL, REHYDROGEL), aluminum phosphate gel, aluminum hydroxyphosphate (ADJU-PHOS), and calcium phosphate.

In another embodiment, hEbola-associated antigen is administered with an immunostimulatory adjuvant. Such class of adjuvants include, but are not limited to, cytokines (e.g., interleukin-2, interleukin-7, interleukin-12, granulocyte-macrophage colony stimulating factor (GM-CSF), interferon-γ, interleukin-1β (IL-1β), and IL-1β peptide or Sevelo Peptide), cytokine-containing liposomes, triterpenoid glycosides or saponins (e.g., Quilla and QS-21), also sold under the trademark STIMULON, ISOCREPE, Muramyl Dipeptide (MDP) derivatives, such as N-acetyl-muramyl-L-threonyl-D-isoglutaminyl (Threonyl-MDP), sold under the trademark THERMUTIDE), GMP, N-acetyl-nor-muramyl-L-alanyl-D-isoglutaminyl, N-acetyl-muramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'2'-dipalmityl-sn-glycero-3-hydroxy phosphatidyl)-ethylamine, muramyl tripeptide phosphatidylethanolamine (MTP-PE), unmethylated CpG dinucleotides and oligonucleotides, such as bacterial DNA and fragments thereof, LPS, monophosphoryl lipid A (3D-MLA sold under the trademark MPL), and polysphazenes.

In another embodiment, the adjuvant is used is a particular adjuvant, including, but not limited to, emulsions, such as Freund’s Complete Adjuvant, Freund’s Incomplete Adjuvant, squalene or squalane oil-in-water adjuvant formulations, such as SAF and MF59, e.g., prepared with block-copolymers, such as L-121 (polyoxypropylene/polyoxyethylene) sold under the trademark PLURONIC L-121, Liposomes, Virosomes, carbonates, and immune stimulating complex, which is sold under the trademark ISCOM.

In another embodiment, a microparticulate adjuvant is used. Microparticulate adjuvants include, but are not limited to, biodegradable and biocompatible polymers, homo- and copolymers of lactic acid (PLA) and glycolic acid (PGA), poly(lactide-co-glycolides) (PLGA) microspheres, polymers that self-assemble into particulates (poloxamer particles), soluble polymers (polysphazenes), and virus-like particles (VLPs) such as recombinant protein particulates, e.g., hepatitis B surface antigen (HbsAg).

Yet another class of adjuvants that are optionally used include mucosal adjuvants, including but not limited to heat-labile enterotoxin from Escherichia coli (LT), cholera holotoxin (CT) and cholera toxin B Subunit (CTB) from Vibrio cholerae, mutant toxins (e.g., LTK05 and LTR72), microparticles, and polymerized liposomes.

In other embodiments, any of the above classes of adjuvants are optionally used in combination with each other or with other adjuvants. For example, non-limiting examples of combination adjuvant preparations used to administer the hEbola-associated antigens of the invention include liposomes containing immunostimulatory protein, cytokines, T-cell and/or B-cell antigens, or microbes with or without entrapped IL-2 or microparticles containing enterotoxin. Other adjuvants known in the art are also included within the scope of the invention (see Vaccine Design: The Subunit and Adjuvant Approach, Chap. 7, Michael F. Powell and Mark J. Newman (eds.), Plenum Press, New York, 1995, which is incorporated herein in its entirety).

The effectiveness of an adjuvant is illustratively determined by measuring the induction of antibodies directed against an immunogenic polypeptide containing a hEbola polypeptide epitope, the antibodies resulting from administration of this polypeptide in vaccines which are also comprised of the various adjuvants.

The polypeptides are optionally formulated into the vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid additional salts (formed with free amino groups of the peptide) and which are formed with inorganic acids, such as, for example, hydrochloric or phosphonic acids, or organic acids such as acetic, oxalic, tartaric, maleic, and the like. Salts formed with free carboxyl groups are optionally derived from inorganic bases, such as, for example, sodium potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, proline and the like.

The vaccines of the invention are preferably multivalent or univalent. Multivalent vaccines are made from recombinant embodiments that direct the expression of more than one antigen.

Many methods are operable herein to introduce the vaccine formulations of the invention; these include but are not limited to oral, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal routes, and via scarification (scratching through the top layers of skin, e.g., using a bifurcated needle). The patient to which the vaccine is administered is preferably a mammal, most preferably a human, but is also optionally a non-human animal including but not limited to lower primates, cows, horses, sheep, pigs, fowl (e.g., chickens), goats, cats, dogs, hamsters, mice and rats.

Preparation of Antibodies

Antibodies that specifically recognize a polypeptide of the invention, such as, but not limited to, polypeptides including the sequence of SEQ ID Nos: 2-9, 59, or 11-19 and other polypeptides as described herein, or hEbola epitope or antigen-binding fragments thereof are used in a preferred embodiment for detecting, screening, and isolating the polypeptide of the invention or fragments thereof, or similar sequences that might encode similar enzymes from other organisms. For example, in one specific embodiment, an antibody which immunospecifically binds hEbola epitope, or a fragment thereof, is used for various in vitro detection assays, including enzyme-linked immunosorbent assays (ELISA), radioimmunoassays, western blot, etc., for the detection of a polypeptide of the invention or, preferably, hEbola, in samples, for example, a biological material, including cells, cell culture media (e.g., bacterial cell culture media, mammalian cell culture media, insect cell culture media, yeast cell
culture media, etc.), blood, plasma, serum, tissues, sputum, nasopharyngeal aspirates, etc.

[0173] Antibodies specific for a polypeptide of the invention or any epitope of hÉbola are optionally generated by any suitable method known in the art. Polyclonal antibodies to an antigen of interest, for example, the hÉbola virus from Deposit Accession No. 200706291, or including a nucleotide sequence of SEQ ID NOs: 1 or 10, are optionally produced by various procedures well known in the art. For example, an antigen is optionally administered to various host animals including, but not limited to, rabbits, mice, rats, etc., to induce the production of antisera containing polyclonal antibodies specific for the antigen. Various adjuvants are optionally used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete) adjuvant, mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, phosphoric polyols, polyoxins, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful adjuvants for humans such as BCG (Bacille Calmette-Guérin) and Corynebacterium parvum. Such adjuvants are also well known in the art.

[0174] Monoclonal antibodies are optionally prepared using a wide variety of techniques known in the art including the use of hybridomas, recombinant, and phage display technologies, or a combination thereof. In one example, monoclonal antibodies are produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., Antibodies: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Monoclonal Antibodies and T-Cell Hybridomas, pp. 563-681 (Elsevier, N.Y., 1981) (both of which are incorporated by reference in their entireties). The term “monoclonal antibody” as used herein is not limited to antibodies produced through hybridoma technology. The term “monoclonal antibody” refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

[0175] Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art. In a non-limiting example, mice are immunized with an antigen of interest or a cell expressing such an antigen. Once an immune response is detected, e.g., antibodies are detected in the mice sera, the mouse spleen is harvested and spleenocytes isolated. The spleenocytes are then fused by well known techniques to any suitable myeloma cells. Hybridomas are selected and cloned by limiting dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding the antigen. Ascites fluid, which generally contains high levels of antibodies, is optionally generated by inoculating mice intraperitoneally with positive hybridoma clones.

[0176] Antibody fragments which recognize specific epitopes are optionally generated by known techniques. For example, Fab and F(ab)2 fragments are illustratively produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab)2 fragments). F(ab)2 fragments preferably contain the complete light chain, and the variable region, the CH1 region and the hinge region of the heavy chain.

[0177] The antibodies of the invention or fragments thereof are optionally produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

[0178] The nucleotide sequence encoding an antibody is obtained from any information available to those skilled in the art (i.e., from Genbank, the literature, or by routine cloning and sequence analysis). If a clone containing a nucleic acid encoding a particular antibody or an epitope-binding fragment thereof is not available, but the sequence of the antibody molecule or epitope-binding fragment thereof is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+RNA, isolated from any tissue or cells expressing the antibody, such as hybidoma cells selected to express an antibody) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, e.g., a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR are optionally then cloned into replicable cloning vectors using any method known in the art.

[0179] Once the nucleotide sequence of the antibody is determined, the nucleotide sequence of the antibody is optionally manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook et al., supra; and Ausubel et al., eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY, which are both incorporated by reference herein in their entireties). To generate antibodies having a different amino acid sequence by, for example, introducing amino acid substitutions, deletions, and/or insertions into the epitope-binding domain regions of the antibodies or any portion of antibodies may enhance or reduce biological activities of the antibodies.

[0180] Recombinant expression of an antibody requires construction of an expression vector containing a nucleotide sequence that encodes the antibody. Once a nucleotide sequence encoding an antibody molecule or a heavy or light chain of an antibody, or portion thereof has been obtained, the vector for the production of the antibody molecule is optionally produced by recombinant DNA technology using techniques known in the art as discussed in the previous sections. Methods which are known to those skilled in the art are optionally used to construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. The nucleotide sequence encoding the heavy-chain variable region, light-chain variable region, both the heavy-chain and light-chain variable regions, an epitope-binding fragment of the heavy- and/or light-chain variable region, or one or more complementarity determining regions (CDRs) of an antibody are optionally cloned into such a vector for expression. Thus, prepared expression vector is optionally then introduced into appropriate host cells for the expression of the antibody. Accordingly, the invention includes host cells containing a polynucleotide encoding an antibody specific for the polypeptides of the invention or fragments thereof.

[0181] The host cell is optionally co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector
encoding a light chain derived polypeptide. The two vectors illustratively contain identical selectable markers which enable equal expression of heavy and light chain polypeptides or different selectable markers to ensure maintenance of both plasmids. Alternatively, a single vector is optionally used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, Nature, 322:52, 1986; and Kohler, Proc. Natl. Acad. Sci. USA, 77:2 197, 1980). The coding sequences for the heavy and light chains optionally include cDNA or genomic DNA.

[0182] In another embodiment, antibodies are generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage is utilized to display antigen binding domains, such as Fab and Fv or disulfide-bond stabilized Fv, expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest is optionally selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phages used in these methods are typically filamentous phage, including fd and M13. The antigen binding domains are expressed as a recombinantly fused protein to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the immunoglobulins, or fragments thereof, of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods, 182:41-50, 1995; Ames et al., J. Immunol. Methods, 184:177-186, 1995; Kettleborough et al., Eur. J. Immunol., 24:952-958, 1994; Persic et al., Gene, 187:9-18, 1997; Barton et al., Advances in Immunology, 57:191-280, 1994; PCT application No. GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Pat. Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,906; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

[0183] As described in the above references, after phage selection, the antibody coding regions from the phage is optionally isolated and used to generate whole antibodies, including human antibodies, or any other desired fragments, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')2 fragments are optionally employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques, 12(6):864-869, 1992; and Sawai et al., AJR, 134:26-34, 1995; and Better et al., Science, 240:1041-1043, 1988 (each of which is incorporated by reference in its entirety). Examples of techniques operable to produce single-chain Fv's and antibodies include those described in U.S. Pat. Nos. 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology, 203:46-88, 1991; Shu et al., PNAS, 90:7995-7999, 1993; and Skerra et al., Science, 240:1038-1040, 1988.

[0184] Once an antibody molecule of the invention has been produced by any methods described above, or otherwise known in the art, it is then optionally purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (e.g., ion exchange, affinity, particularly by affinity for the specific antigen after Protein A or Protein G purification, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique(s) for the purification of proteins. Further, the antibodies of the present invention or fragments thereof are optionally fused to heterologous polypeptide sequences described herein or otherwise known in the art to facilitate purification. Illustrative examples include 6×His tag, FLAG tag, biotin, avidin, or other system.

[0185] For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it is preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a constant region derived from a human immunoglobulin. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science, 229:1202, 1985; Oi et al., BioTechniques, 4:214 1986; Gillies et al., J. Immunol. Methods, 125:191-202, 1989; U.S. Pat. Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species that bind the desired antigen having one or more complementarity determining regions (CDRs) from the non-human species and framework regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. See, e.g., Queen et al., U.S. Pat. No. 5,585,089; Riechmann et al., Nature, 332:323, 1988, which are incorporated herein by reference in their entirety. Antibodies are humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Pat. Nos. 5,225,339; 5,530,101 and 5,583,089), “shuffling” or “resurfacing” (EP 509,978; 502,106; EP 519,596; Padlan, Molecular Immunology, 28(4/5): 489-498, 1991; Studnicka et al., Protein Engineering, 7(6): 805-814, 1994; Roguska et al., Proc Natl Acad. Sci. USA, 91:969-973, 1994), and chain shuffling (U.S. Pat. No. 5,565,332), all of which are hereby incorporated by reference in their entirety.

[0186] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies are made by a variety of methods known in the art illustratively including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See U.S. Pat. Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645; WO 98/50433; WO 98/24893; WO 98/16654; WO 96/34096; WO 96/33735; and WO 91/10741, each of which is incorporated herein by reference in its entirety.
detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 5 98 787; U.S. Pat. Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entireties. In addition, companies such as Abgenix, Inc. (Fremont, Calif.), Medarex (NJ) GenPharm (San Jose, Calif.) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

[0188] Completely human antibodies which recognize a selected epitope are optionally generated using a technique referred to as “guided selection.” In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jesper et al., Bio/technology, 12:899-903, 1988).

[0189] Antibodies fused or conjugated to heterologous polypeptides are optionally used in in vitro immunosays and in purification methods (e.g., affinity chromatography) known in the art. See e.g., PCT publication No. WO 93/21232; EP 439,805; Naramura et al., Immunol. Lett., 39:91-99, 1994; U.S. Pat. No. 5,474,981; Gillies et al., PNAS, 89:1428-1432, 1992; and Fell et al., J. Immunol., 146:2446-2452, 1991, which are incorporated herein by reference in their entireties.

[0190] Antibodies may also be illustratively attached to solid supports, which are particularly useful for immunosays or purification of the polypeptides of the invention or fragments, derivatives, analogs, or variants thereof, or similar molecules having the similar enzymatic activities as the polypeptide of the invention. Such solid supports include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypethylene.

Pharmaceutical Compositions and Kits

[0191] The present invention encompasses pharmaceutical compositions including antiviral agents of the present invention. In a specific embodiment, the antiviral agent is preferably an antibody which immunospecifically binds and neutralizes the hEbola virus or variants thereof, or any proteins derived therefrom. In another specific embodiment, the antiviral agent is a polypeptide or nucleic acid molecule of the invention. The pharmaceutical compositions have utility as an antiviral prophylactic agent are illustratively administered to a subject where the subject has been exposed or is expected to be exposed to a virus.

[0192] Various delivery systems are known and operable to administer the pharmaceutical composition of the invention, illustratively, encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the mutant viruses, and receptor mediated endocytosis (see, e.g., Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432). Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epideral, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and optionally administered together with other biologically active agents. Administration is systemic or local. In a preferred embodiment, it is desirable to introduce the pharmaceutical compositions of the invention into the lungs by any suitable route. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

[0193] In a specific embodiment, it is desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment. This administration may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, by means of nasal spray, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection at the site (or former site) infected tissues.

[0194] In another embodiment, the pharmaceutical composition is delivered in a vesicle, in particular a liposome (see Langer, 1990, Science 249:1527-1533; Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; see generally ibid.).


[0196] Other controlled release systems are discussed in the review by Langer (Science 249:1527-1533 (1990)) the contents of which are incorporated herein by reference.

[0197] The pharmaceutical compositions of the present invention illustratively include a therapeutically effective amount of a live attenuated, inactivated or killed West African hEbola virus, or recombinant or chimeric hEbola virus, and a pharmaceutically acceptable carrier. In a specific embodiment, the term “pharmaceutically acceptable” means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term “carrier” refers to a diluent, adjuvant, excipient, or vehicle with which the pharmaceutical composition is administered. Such pharmaceutical carriers are illustratively sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions are optionally
employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, tallow, sodium chloride, dried skim milk, glycerol, propylene glycol, water, ethanol and the like. The composition, if desired, also contains wetting or emulsifying agents, or pH buffering agents. These compositions optionally take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained release formulations and the like. The composition is optionally formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation illustratively includes standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in “Remington’s Pharmaceutical Sciences” by E. W. Martin. The formulation should suit the mode of administration.

[0198] In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. The composition also includes an optional solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water-free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline is optionally provided so that the ingredients may be mixed prior to administration.

[0199] The pharmaceutical compositions of the invention are illustratively formulated as neutral or salt forms. Pharmacologically acceptable salts illustratively include those formed with free amino groups such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isoproplamine, triethylenediamine, 2 ethylamino ethanol, histidine, proprazine, etc.

[0200] The amount of the pharmaceutical composition of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, in vitro assays are optionally employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient’s circumstances. However, suitable dosage ranges for intravenous administration are generally about 20 to 500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose response curves derived from in vitro or animal model test systems.

[0201] Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

[0202] The invention also provides a pharmaceutical pack or kit including one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) is a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In a preferred embodiment, the kit contains an antiviral agent of the invention, e.g., an antibody specific for the polypeptides encoded by a nucleotide sequence of SEQ ID NOs: 1 or 10, or as shown in SEQ ID NOs: 2-9, 59, or 11-19, or any hEbo-la epitope, or a polypeptide or protein of the present invention, or a nucleic acid molecule of the invention, alone or in combination with adjuvants, antivirals, antibiotics, analgesics, bronchodilators, or other pharmaceutically acceptable excipients.

[0203] The present invention further encompasses kits including a container containing a pharmaceutical composition of the present invention and instructions for use.

Detection Assays

[0204] The present invention provides a method for detecting an antibody, which immunospecifically binds to the hEbo-la virus, in a biological sample, including for example blood, serum, plasma, saliva, urine, feces, etc., from a patient suffering from hEbo-la infection, and/or hemorrhagic fever. In a specific embodiment, the method including contacting the sample with the hEbo-la virus, for example, of Deposit Accession No. 200706291, or having a genomic nucleic acid sequence of SEQ ID NOs: 1 or 10, directly immobilized on a substrate and detecting the virus-bound antibody directly or indirectly by a labeled heterologous anti-isotype antibody. In another specific embodiment, the sample is contacted with a host cell which is infected by the hEbo-la virus, for example, of Deposit Accession No. 200706291, or having a genomic nucleic acid sequence of SEQ ID NOs: 1 or 10, and the bound antibody is optionally detected by immunofluorescent assay.

[0205] An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid of the invention in a biological sample involves obtaining a biological sample from various sources and contacting the sample with a compound or an agent capable of detecting an epitope or nucleic acid (e.g., mRNA, genomic DNA) of the hEbo-la virus such that the presence of the hEbo-la virus is detected in the sample. A preferred agent for detecting hEbo-la mRNA or genomic RNA of the invention is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic RNA encoding a polypeptide of the invention. The nucleic acid probe is, for example, a nucleic acid molecule including the nucleotide sequence of SEQ ID NOs: 1 or 10, a complement thereof, or a portion thereof, such as an oligonucleotide of at least 15, 20, 25, 30, 50, 100, 250, 500, 750, 1000 or more contiguous nucleotides in length and sufficient to specifically hybridize under stringent conditions to a hEbo-la mRNA or genomic RNA.

[0206] As used herein, the term “stringent conditions” describes conditions for hybridization and washing under which nucleotide sequences having at least 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity to each other typically remain hybridized to
each other. Such hybridization conditions are described in, for example but not limited to, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1 6.3.6; Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., N.Y. (1986), pp. 75, 78, and 84, 87; and Molecular Cloning, Cold Spring Harbor Laboratory, N.Y. (1982), pp. 387, 389, and are well known to those skilled in the art. A preferred, non-limiting example of stringent hybridization conditions is hybridization in 6x sodium chloride/sodium citrate (SSC), 0.5% SDS at about 68°C followed by one or more washes in 2xSSC, 0.5% SDS at room temperature. Another preferred, non-limiting example of stringent hybridization conditions is hybridization in 6xSSC at about 45°C followed by one or more washes in 0.2xSSC, 0.1% SDS at 50 to 65°C.

[0207] A nucleic acid probe, polynucleotide, oligonucleotide, or other nucleic acid is preferably purified. An “isolated” or “purified” nucleotide sequence is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the nucleotide is derived, or is substantially free of chemical precursors or other chemicals when chemically synthesized. The language “substantially free of cellular material” includes preparations of a nucleotide/oligonucleotide in which the nucleotide/oligonucleotide is separated from cellular components of the cells from which it is isolated or produced. Thus, a nucleotide/oligonucleotide that is substantially free of cellular material includes preparations of the nucleotide having less than about 30%, 20%, 10%, 5%, 2.5%, or 1%, (by dry weight) of contaminating material. When nucleotide/oligonucleotide is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, i.e., it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. Accordingly, such preparations of the nucleotide/oligonucleotide have less than about 30%, 20%, 10%, or 5% (by dry weight) of chemical precursors or compounds other than the nucleotide/oligonucleotide of interest. In a preferred embodiment of the present invention, the nucleotide/oligonucleotide is isolated or purified.

[0208] In another preferred specific embodiment, the presence of hEbola virus is detected in the sample by a reverse transcription polymerase chain reaction (RT-PCR) using the primers that are constructed based on a partial nucleotide sequence of the genome of hEbola virus, for example, that of Deposit Accession No. 200706291, or having a genomic nucleic acid sequence of SEQ ID NOs: 1 or 10. In a non-limiting specific embodiment, preferred primers to be used in a RT-PCR method are the primers are described in detail herein.

[0209] In more preferred specific embodiment, the present invention provides a real-time quantitative PCR assay to detect the presence of hEbola virus in a biological sample by subjecting the cDNA obtained by reverse transcription of the extracted total RNA from the sample to PCR reactions using the specific primers described in detail herein, and a fluorescence dye, such as SYBR® Green I, which fluoresces when bound nonspecifically to double-stranded DNA. The fluorescence signals from these reactions are captured at the end of extension steps as PCR product is generated over a range of the thermal cycles, thereby allowing the quantitative determination of the viral load in the sample based on an amplification plot.

[0210] A preferred agent for detecting hEbola is an antibody that specifically binds a polypeptide of the invention or any hEbola epitope, preferably an antibody with a detectable label. Antibodies are illustratively polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')2) is operable herein.

[0211] The term “labeled”, with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, optionally via a linker, as well as indirect labeling of the probe or antibody by reaction with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it is detectable with fluorescently labeled streptavidin. The detection method of the invention is optionally used to detect mRNA, protein (or any epitope), or genomic RNA in a sample in vitro as well as in vivo. Exemplary in vitro techniques for detection of mRNA include northern hybridizations, situ hybridizations, RT-PCR, and RNase protection. In vitro techniques for detection of an epitope of hEbola illustratively include enzyme linked immunosorbent assays (ELISAs), western blots, immunoprecipitations and immunofluorescence. In vivo techniques for detection of genomic RNA include northern hybridizations, RI-PCR, and RNase protection. Furthermore, in vivo techniques for detection of hEbola include introducing into a subject organism a labeled antibody directed against the polypeptide. In one embodiment, the antibody is labeled with a radioactive marker whose presence and location in the subject organism is detected by standard imaging techniques, including autoradiography.

[0212] In a specific embodiment, the methods further involve obtaining a control sample from a control subject, contacting the control sample with a compound or agent capable of detecting hEbola, e.g., a polypeptide of the invention or mRNA or genomic RNA encoding a polypeptide of the invention, such that the presence of hEbola or the polypeptide or mRNA or genomic RNA encoding the polypeptide is detected in the sample, and comparing the absence of hEbola or the polypeptide or mRNA or genomic RNA encoding the polypeptide in the control sample with the presence of hEbola, or the polypeptide or mRNA or genomic DNA encoding the polypeptide in the test sample.

[0213] The invention also encompasses kits for detecting the presence of hEbola or a polypeptide or nucleic acid of the invention in a test sample. The kit illustratively includes a labeled compound or agent capable of detecting hEbola or the polypeptide or a nucleic acid molecule encoding the polypeptide in a test sample and, in certain embodiments, a means for determining the amount of the polypeptide or mRNA in the sample (e.g., an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits optionally include instructions for use.

[0214] For antibody-based kits, the kit illustratively includes: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide of the invention or hEbola epitope; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is preferably conjugated to a detectable agent.

[0215] For oligonucleotide-based kits, the kit illustratively includes: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence
encoding a polypeptide of the invention or to a sequence within the hEbola genome; or (2) a pair of primers useful for amplifying a nucleic acid molecule containing an hEbola sequence. The kit optionally includes a buffering agent, a preservative, or a protein stabilizing agent. The kit optionally includes components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit optionally contains a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for use.

Screening Assays to Identify Antiviral Agents

[0216] The invention provides methods for the identification of a compound that inhibits the ability of hEbola virus to infect a host or a host cell. In certain embodiments, the invention provides methods for the identification of a compound that reduces the ability of hEbola virus to replicate in a host or a host cell. Any technique well known to the skilled artisan is illustratively used to screen for a compound useful to abolish or reduce the ability of hEbola virus to infect a host and/or to replicate in a host or a host cell.

[0217] In certain embodiments, the invention provides methods for the identification of a compound that inhibits the ability of hEbola virus to replicate in a mammal or a mammalian cell. More specifically, the invention provides methods for the identification of a compound that inhibits the ability of hEbola virus to infect a mammal or a mammalian cell. In certain embodiments, the invention provides methods for the identification of a compound that inhibits the ability of hEbola virus to replicate in a mammalian cell. In a specific embodiment, the mammalian cell is a human cell.

[0218] In another embodiment, a cell is contacted with a test compound and infected with the hEbola virus. In certain embodiments, a control culture is infected with the hEbola virus in the absence of a test compound. The cell is optionally contacted with a test compound before, concurrently with, or subsequent to the infection with the hEbola virus. In a specific embodiment, the cell is a mammalian cell. In an even more specific embodiment, the cell is a human cell. In certain embodiments, the cell is incubated with the test compound for at least 1 minute, at least 5 minutes, at least 15 minutes, at least 30 minutes, at least 1 hour, at least 2 hours, at least 5 hours, at least 12 hours, or at least 1 day. The titer of the virus is optionally measured at any time during the assay. In certain embodiments, a time course of viral growth in the culture is determined. If the viral growth is inhibited or reduced in the presence of the test compound, the test compound is identified as being effective in inhibiting or reducing the growth or infection of the hEbola virus. In a specific embodiment, the compound that inhibits or reduces the growth of the hEbola virus is tested for its ability to inhibit or reduce the growth rate of other viruses to test its specificity for the hEbola virus.

[0219] In one embodiment, a test compound is administered to a model animal and the model animal is infected with the hEbola virus. In certain embodiments, a control model animal is infected with the hEbola virus without the administration of a test compound. The test compound is optionally administered before, concurrently with, or subsequent to the infection with the hEbola virus. In a specific embodiment, the model animal is a mammal. In an even more specific embodiment, the model animal is, but is not limited to, a cotton rat, a mouse, or a monkey. The titer of the virus in the model animal is optionally measured at any time during the assay. In certain embodiments, a time course of viral growth in the culture is determined. If the viral growth is inhibited or reduced in the presence of the test compound, the test compound is identified as being effective in inhibiting or reducing the growth or infection of the hEbola virus. In a specific embodiment, the compound that inhibits or reduces the growth of the hEbola in the model animal is tested for its ability to inhibit or reduce the growth rate of other viruses to test its specificity for the hEbola virus.

[0220] According to the method of the invention, a human or an animal is optionally treated for for EboBun or EboIC, other viral infection or bacterial infection by administering an effective amount of an inventive therapeutic composition. Preferably, a vaccine is administered prophylactically. An “effective amount” is an amount that will induce an immune response in a subject. Illustratively, an effective amount of the compositions of this invention ranges from nanogram/kg to milligram/kg amounts for young children and adults. Equivalent dosages for lighter or heavier body weights can readily be determined. The dose should be adjusted to suit the individual to whom the composition is administered and will vary with age, weight and metabolism of the individual. The exact amount of the composition required will vary from subject to subject, depending on the species, age, weight and general condition of the subject, the particular peptide or polypeptide used, its mode of administration and the like. An appropriate amount can be determined by one of ordinary skill in the art using only routine experimentation given the teachings herein. One skilled in the art will realize that dosages are best optimized by the practicing physician or veterinarian and methods for determining dose amounts and regimens and preparing dosage forms are described, for example, in Remington’s Pharmaceutical Sciences, (Martin, E. W., ed., latest edition), Mack Publishing Co., Easton, Pa. Preferably, a single administration is operable to induce an immune response.


[0222] Embodiments of inventive compositions and methods are illustrated in the following detailed examples. These examples are provided for illustrative purposes and are not considered limitations on the scope of inventive compositions and methods.

EXAMPLES

Example 1

Newly Discovered Ebola Virus Associated with Hemorrhagic Fever Outbreak in Bundibugyo, Uganda

[0223] In late November 2007 HF cases were reported in the townships of Bundibugyo and Kikyo in Bundibugyo Dis-
district, Western Uganda (FIG. 1A). These samples were assayed as described by Towner, JS, et al., PLoS Pathog. 2008 November; 4(11): e1000212, the contents of which are incorporated herein by reference for methods, results, reagents, and all other aspects of the publication. A total of 29 blood samples were initially collected from suspect cases and showed evidence of acute ebolavirus infection in eight specimens using a broadly reactive ebolavirus antigen capture assay known to cross-react with the different ebolavirus species and an IgM capture assay based on Zaire ebolavirus reagents (Table 1). These specimens were negative when initially tested with highly sensitive real-time RT-PCR assays specific for all known Zaire and Sudan ebolaviruses and marburgviruses. However, further evidence of acute ebolavirus infection was obtained using a traditionally less sensitive (relative to the real-time RT-PCR assays) but more broadly reactive filovirus L gene-specific RT-PCR assay (1 specimen) (Table 1). Sequence analysis of the PCR fragment (400 bp of the virus L gene) revealed the reason for the initial failure of the real-time RT-PCR assays, as the sequence was distinct from that of the 4 known species of ebolavirus, although distantly related to Côte d'Ivoire ebolavirus. In total, 9 of 29 specimens showed evidence of ebolavirus infection, and all tests were negative for marburgvirus (data not shown).

[0224] Approximately 70% of the virus genome was rapidly sequenced from total RNA extracted from a patient serum (#200706291) using a newly established metagenomic pyrosequencing method (454 Life Sciences) which involves successive rounds of random DNA amplification. Using the newly derived draft sequence, a real-time RT-PCR assay specific for the NP gene of this virus was quickly developed and evaluated. The assay was shown to have excellent sensitivity (Table 1), finding positive all the initial six samples that tested positive by either virus antigen capture (five specimens) or virus isolation assays (four specimens). The antigen-capture, IgM, IgG and newly designed real-time PCR assays were quickly transferred to the Uganda Virus Research Institute during the course of the outbreak to facilitate rapid identification and isolation of ebolavirus cases in the affected area for efficient control of the outbreak. The outbreak continued through late December 2007, and resulted in 149 suspected cases and 37 deaths.

[0225] Table 1. Ebolavirus diagnostic results of initial 29 specimens obtained from Bundibugyo District with numerical specimen numbers assigned. RT-PCR refers to results obtained from conventional PCR using the broadly reactive Filo A/B primers. Ag, IgM, and IgG refer to results from ELISA-based assays with Zaire ebolavirus reagents while virus isolation refers to culture attempts on Vero E6 cells. Q-RT-PCR refers to results obtained using the optimized Bundibugyo ebolavirus specific real-time RT-PCR assay with cycle threshold (Ct) values of positive (Pos) samples indicated in the far right column. * Specimen #200706291 is the clinical sample from which prototype isolate #811250 was obtained.

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>RT-PCR</th>
<th>Ag</th>
<th>IgM</th>
<th>IgG</th>
<th>Virus Isolation</th>
<th>Q-RT-PCR</th>
<th>Ct</th>
</tr>
</thead>
<tbody>
<tr>
<td>200706288</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>40</td>
</tr>
<tr>
<td>200706289</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>40</td>
</tr>
<tr>
<td>200706290</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
<td>40</td>
</tr>
<tr>
<td>200706291*</td>
<td>Pos</td>
<td>Pos</td>
<td>neg</td>
<td>neg</td>
<td>Pos</td>
<td>Pos</td>
<td>23.64</td>
</tr>
</tbody>
</table>

[0226] The entire genome sequence of this virus was completed using a classic primer walking sequencing approach on RNA. The complete genome of the Ebolavirus was not available, so it too was derived by a similar combination of random primed pyrosequencing and primer walking approaches. Acquisition of these sequences allowed for the first time the phylogenetic analysis of the complete genomes of representatives of all known species of Ebola and Marburg viruses. The analysis revealed that the newly discovered virus differed from the four existing ebolavirus species (FIG. 1), with approximately 32% nucleotide difference from even the closest relative, EboIC (Table 2). Similar complete genome divergence (35-45%) is seen between the previously characterized ebolavirus species.

[0227] Table 2. Identity matrix based on comparisons of full-length genome sequences of Zaire ebolaviruses 1976 (Genbank accession number NC_002549) and 1995 (Genbank accession number AY354458), Sudan ebolavirus 2000 (Genbank accession number NC_006452), Cote d'Ivoire ebolavirus 1994 (SEQ ID NO. 10), Reston ebolavirus 1989 (Genbank accession number NC_004161), and Bundibugyo ebolavirus 2007 (SEQ ID NO. 1).

<table>
<thead>
<tr>
<th>Zaire '76</th>
<th>'95</th>
<th>Sudan '00</th>
<th>EboIC '94</th>
<th>EboBun '07</th>
<th>Reston '89</th>
</tr>
</thead>
<tbody>
<tr>
<td>.988</td>
<td>.577</td>
<td>.630</td>
<td>.632</td>
<td>.581</td>
<td></td>
</tr>
<tr>
<td>.577</td>
<td>.631</td>
<td>.633</td>
<td>.581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.577</td>
<td>.577</td>
<td>.683</td>
<td>.575</td>
<td>.576</td>
<td></td>
</tr>
</tbody>
</table>
and makes possible the development of clinical, diagnostic and research tools directed to human hEbola infection.

Material and Methods


[0230] Several diagnostic techniques were used for each sample: (i) antigen capture, lgG, and IgM assays were performed as previously described[11]; (ii) virus isolation attempts were performed on Vero E6 cells[12] and monitored for 14 days; (iii) RNA was extracted and tested for Zaire[13] and Sudan ebolavirus and marburgvirus[14] using real-time quantitative RT-PCR assays designed to detect all known species of each respective virus the primers/probe for the Sudan ebolavirus assay were EboSuDBMG 1(+) 5'-GCC ATG GTC TCA GTG TG-3' (SEQ ID NO: 21); EboSuDBMG 1 (-) 5'-GTT IAC ATT GGC CAA CAA TCC A-3' (SEQ ID NO: 22) and Ebola Sudan BMG Probe 5'FAM-AC GTC GAT CAT TCT TCT CTC GGA BHQ1 (SEQ ID NO: 23); (iv) the conventional RT-PCR was performed with the file A/B primer set as previously described[15] using Superscript III (Invitrogen) according to the manufacturer's instructions. The specimen 20076291 was selected as the reference sample for further sequence analysis.

[0231] Genome Sequencing.

[0232] Pyrosequencing was carried out utilizing the approach developed by 454 Life Sciences, and the method described by Cox-Foster et al.[3]. Subsequent virus whole genome primer walking was performed as previously described[16] but using the primers specific for Bundibugyo ebolavirus RT-PCR amplification. In total, the entire virus genome was amplified in six overlapping RT-PCR fragments (all primers listed 5' to 3'): fragment A (predicted size 2.7 kb) was amplified using forward-GTGAGCGAAAACATCACPCTCTGG (SEQ ID NO: 24) with reverse-CACTCAAGTTGCTAGAGATCC (SEQ ID NO: 25); fragment B (predicted size 3.0 kb) was amplified using forward-CCACAAAACACACCTGAGCTAAGT (SEQ ID NO: 26) with reverse-AGGTCGCGTATACCTTCTAC (SEQ ID NO: 27); fragment C (predicted size 3.5 kb) was amplified using forward-GATGGTTTATGATCTTTCCG (SEQ ID NO: 28) with reverse-GTCTTTGARTACATCAATG (SEQ ID NO: 29); fragment D (predicted size 3.1 kb) was amplified using forward-CCACAGACCACAGAGAC (SEQ ID NO: 30) with reverse-CTATCGGCAATGGACATTT (SEQ ID NO: 31); fragment E (predicted size 3.4 kb) was amplified using forward-GGCGTCTTAGAGAGCACAC (SEQ ID NO: 32) with reverse-CACATTAATGTTTCTACAGTGCAAG (SEQ ID NO: 33) and fragment F (predicted size 3.5 kb) was amplified using forward-CTCATGTATTTAGACAGGACTA (SEQ ID NO: 34) with reverse-GGTTAGTAGAGGAAT (SEQ ID NO: 35).

[0233] The exact 5' and 3' ends of Bundibugyo ebolavirus were determined by 3' RACE from virus RNA extracted from virus infected Vero E6 cell monolayers using TriPure isolation reagent. RNAs were then polyadenylated in vitro using A-Plus poly(A) polymerase tailing kit (Epicerent Biotechnologies) following the manufacturer's instructions and then purified using an RNeasy kit (Qiagen) following standard protocols. Ten microliters of in vitro polyadenylated RNA were added as template in RT-PCR reactions, using SuperScript III One-Step RT-PCR system with Platinum Taq High Fidelity (Invitrogen) following the manufacturer's protocol. Two parallel RT-PCR reactions using the oligo(dT)-containing 3' RACE-AP primer (Invitrogen) mixed with 1 of 2 viral specific primers, Ebo-U 692(-) ACAAAAAAGCTAICTGACTAT (SEQ ID NO: 36) and Ebo-V18269(+) CTCAAGGCAAATTAATGG (SEQ ID NO: 37), generated ~700 nt long fragments containing the 3' ends of either genomic and antigenic RNAs. The resulting RT-PCR products were analyzed by agarose electrophoresis, and DNA bands of the correct sizes were purified using QIAquick Gel Extraction Kit (Qiagen) and sequenced using standard protocols (ABI).

[0234] The nucleotide sequence of the Côte d'Ivoire ebolavirus (EboC) isolate RNA was initially determined using the exact same pyrosequencing strategy as that used for Bundibugyo ebolavirus described above. This method generated sequence for approximately 70% of the entire genome. This draft sequence was then used to design a whole genome primer walking strategy for filling any gaps and confirming the initial sequence. The following Côte d'Ivoire ebolavirus-specific primers were used to generate RT-PCR fragments, designated A-F, as follows: Fragment A (predicted size 3.0 kb) was amplified using forward-GTGTTGCGGAATACTGTAAGGGAAG (SEQ ID NO: 38) and reverse-GTCTTGTGCAATGTTGAAGG (SEQ ID NO: 39); Fragment B (predicted size 3.2 kb) was amplified using forward-CATTGAAACCACACTCCACCG (SEQ ID NO: 40) and reverse-GTTGCTTTAACTCTACTAAGT (SEQ ID NO: 41); Fragment C (predicted size 3.0 kb) was amplified using forward-GGTCTAATGTTTTTCTCGCG (SEQ ID NO: 42) and reverse-CAAGTTTATGTTTGTTGTCCTCAG (SEQ ID NO: 43); Fragment D (predicted size 3.5 kb) was amplified using forward-GTCTGGAATGAGATCACCAG (SEQ ID NO: 44) and reverse-GGCTAGTCTACTAGTCCTGTA (SEQ ID NO: 45); Fragment E (predicted size 4.0 kb) was amplified using forward-GACAAGAAGAGATGGAGGC (SEQ ID NO: 46) and reverse-GTAAATGAGGAGGGTCTTGG (SEQ ID NO: 47); Fragment F (predicted size 2.9 kb) was amplified using forward-CACACAGTTGTGACCAATTTGG (SEQ ID NO: 48) and reverse-CAGACACTAATAGATCTGGTGAAG (SEQ ID NO: 49); Fragment G (predicted size 1.3 kb) was amplified using forward-GGCAACCAAAAAGAAAWRAA (SEQ ID NO: 50) and reverse-GTGCTTCTAGTCTGCTC (SEQ ID NO: 51); and fragment H (predicted size 2.5 kb) was amplified using forward-GTATATTATAGCTGTAAGGCT (SEQ ID NO: 52) and reverse-TGGAGCACCAAAAAARGAACAGGA (SEQ ID NO: 53). A gap in the sequence contig was located between fragments C and D and this was resolved using the following primers to generate a predicted fragment of 1.5 kb: forward-CGGTACACAGATCCAGAAGAAG (SEQ ID NO: 54) and reverse-GTGTAAAAGCTGGTAGATACCTCC (SEQ ID NO: 55). The terminal ~20 nucleotides of the sequence were not experimentally determined but were inferred by comparing with the other known Ebola genome sequences.


[0236] The primers and probe used in the Bundibugyo ebolavirus specific Q-RT-PCR assay were as follows: EboU965 (+): 5'-GAGAAAAAGGCCGCTGTCTGGGAAT-3' (SEQ ID NO: 56); EboU1093(-): 5'-TCCGGTGTATGTAATCGACACTGTGGTT-3' (SEQ ID NO: 57) and EboU989 Prb: 5'-FAM-TTCAAGCACAAATCTCCATGCACCGCA-3' BHQ1 (SEQ ID NO 58). Q-RT-PCR reactions were set up using SuperScript III One-Step Q-RT-PCR (Invitrogen) according to the manufacturer's instructions and run for 40 cycles with a 58° C. annealing temperature.
Phylogenetic Analysis. Modeltest 3.7 was used to examine 56 models of nucleotide substitution to determine the model most appropriate for the data. The General Time Reversible model incorporating invariant sites and a gamma distribution (GTR+I+G) was selected using the Akaike Information Criterion (AIC). Nucleotide frequencies were A=0.3278, C=0.2101, G=0.1832, T=0.2789, the proportion of invariant sites=0.1412, and the gamma shape parameter=1.0593. A maximum likelihood analysis was subsequently performed in PAUP*. 0b10^19 using the GTR+I+G model parameters. Bootstrap support values were used to assess topological support and were calculated based on 1,000 pseudoreplicates.

In addition, a Bayesian phylogenetic analysis was conducted in MrBayes 3.2 with the GTR+I+G model of nucleotide substitution. Two simultaneous analyses, each with four Markov chains, were run for 5,000,000 generations sampling every 100 generations. Prior to termination of the run, the AWTY module was used to assess Markov Chain Monte Carlo convergence to ensure that the length of the analysis was sufficient. Trees generated before the stabilization of the likelihood scores were discarded (burn in=40), and the remaining trees were used to construct a consensus tree. Nodal support was assessed by posterior probability values (>95=statistical support).

Example 2: Immunization against Ebola

To determine the capability of immunogens to elicit an immune response in non-human primates (NHP), 12 cynomolgus macaques, of which 10 are immunized with VSVAG/EboBunGP either orally (OR; n=4), intranasally (IN; n=4) or intramuscularly (IM; n=2) in accordance with all animal control and safety guidelines and essentially as described by Qiu, X, et al., PLoS ONE. 2009; 4(5): e5547. The remaining 2 control animals are vaccinated intramuscularly with VSVAG/MARVGp. VSVAG/MARVGp does not provide heterologous protection against Ebola, therefore these NHPs succumb to Ebola infection. Animals are acclimatized for 14 days prior to infection. Animals are fed and monitored twice daily (pre- and post-infection) and fed commercial monkey Chow, treats and fruit. Husbandry enrichment consists of commercial toys and visual stimulation.

The recombinant VSVAG/EboBun vaccines are synthesized expressing the EboBun glycoprotein (GP) (SEQ ID NO: 9), soluble glycoprotein (sGP) (SEQ ID NO: 4), or nucleoprotein (NP) (SEQ ID NO: 3). Control VSVAG/MARVGp vaccines represent the analogous proteins from Lake victoria marburgvirus (MARV) (strain Musoke). The following results for GP are similar for sGP and NP. Vaccines are generated using VSV (Indiana serotype) as described previously. Garbutt, M, et al., J Virol, 2004; 78(10):5458-5465; Schnell, M J, et al., J Infect Dis, 1999; 179 (Suppl 1):S224-34. An EboBun immunogen peptide pool consisting of 15mers with 11 amino acid overlaps (Sigma-Genosys) spanning the entire sequence of the EboBun immunogens and strain Mayinga 1976 GP are used.

Twelve filovirus naïve cynomolgus monkeys randomized into four groups receive 2 ml of 1×10^7 PFU/ml of vaccine in Dulbecco’s modified Eagle’s medium (DMEM), Animals in the three experimental groups are vaccinated with either: 1) 2 ml orally (OR; n=4); 2) 1 ml dripped into each nostril, intranasally (IN; n=4); or 3) 1 ml each into two sites intramuscularly (IM; n=2). The two controls are injected intramuscularly with 2 ml of 1×10^7 PFU/ml of VSVAG/MARVGp. All animals are challenged intramuscularly 28 days later with 1,000 PFU of EboBun.

Routine examination is conducted on 0, 2, 4, 6, 10, 14 and 21 days post-vaccination, then 0, 3, 6, 10, 14, 19, 20 days, 6 and 9 months after the EboBun challenge. For the examinations animals are anaesthetized by intramuscular injection with 10 mg/kg of ketaset (Ayers). Examinations include haematological analysis, monitoring temperature (rectal), respiration rate, lymph nodes, weight, hydration, discharges and mucous membranes. Also, swabs (throat, oral, nasal, rectal, vaginal) and blood samples are collected (4 ml from femoral vein, 1 ml in EDTA vacutainer tube; 3 ml in serum separator vacutainer tube). Cynomolgus monkey PBMCs are isolated using BD CPT sodium citrate Vacutainers (Becton Dickinson) as per manufacturer’s protocol.

All VSVAG/EboBunGP immunized animals are protected from high dose challenge. These animals show no evidence of clinical illness after vaccination or EboBun challenge. Both control animals demonstrate typical symptoms associated with EboBun infection. Two animals were euthanized on day 4 post-infection due to severe clinical illness. Hematological analysis at each examination date demonstrate increases in the platelet-crit in the OR and IN groups post-challenge, however, no significant changes are observed in any NHPs post-inmunization or in the VSVAG/EboBunGP immunized NHPs post-challenge.

All animals in the three experimental groups are vaccinated with either: 1) 2 ml intramuscularly (IM; n=4); 2) 1 ml dripped into each nostril, intranasally (IN; n=4); or 3) 1 ml each into two sites intramuscularly (IM; n=2). The two controls are injected intramuscularly with 2 ml of 1×10^7 PFU/ml of VSVAG/MARVGp. All animals are challenged intramuscularly 28 days later with 1,000 PFU of EboBun.

Animals in the three experimental groups are vaccinated with either: 1) 2 ml orally (OR; n=4); 2) 1 ml dripped into each nostril, intranasally (IN; n=4); or 3) 1 ml each into two sites intramuscularly (IM; n=2). The two controls are injected intramuscularly with 2 ml of 1×10^7 PFU/ml of VSVAG/MARVGp. All animals are challenged intramuscularly 28 days later with 1,000 PFU of EboBun.

All VSVAG/EboBunGP immunized animals are protected from high dose challenge. These animals show no evidence of clinical illness after vaccination or EboBun challenge. Both control animals demonstrate typical symptoms associated with EboBun infection. Two animals were euthanized on day 4 post-infection due to severe clinical illness. Hematological analysis at each examination date demonstrate increases in the platelet-crit in the OR and IN groups post-challenge, however, no significant changes are observed in any NHPs post-inmunization or in the VSVAG/EboBunGP immunized NHPs post-challenge.


The VSVAG/MARVGp immunized animals do not develop a detectable antibody response to EboBun. In contrast, potent antibody responses are detected in all VSVAG/ EboBunGP immunized animals independent of immunization route. Between days 14 and 21 post-vaccination, all VSVAG/EboBunGP immunized NHPs develop high levels of IgG, IgM, and IgG against EboBunGP. After challenge the IgM titres do not exceed the post-vaccination levels, however, IgG and IgA antibody titres are increased peaking 14 days post-challenge then slowly decreasing before maintaining a relatively high antibody titre up to 9 months.

Neutralization antibodies are detected by a EboBun-GFP flow cytometric neutralization assay in serum collected at days 0 and 21 post-vaccination. Samples are assayed in duplicate for their ability to neutralize an infection with EboBun-GFP in VeroE6 cells. Serially diluted serum samples are incubated with an equal volume of EboBun-GFP in DMEM, at 37°C, 5% CO2 for 1 hr followed by addition of 150 µl per well of a confluent 12 well plate of VeroE6 cells (MOI=0.0005). After 2 hours at 37°C, 5% CO2, 1 ml of DMEM, 2% fetal bovine serum (FBS), 100 µl penicillin, 100 µg/ml streptomycin is added per well and incubated for 5 days. Cells are harvested by removing the culture supernatant, washing with 1 ml PBS, 0.04% EDTA, then adding 800 µl of PBS 0.04% EDTA for 5 minutes at 37°C before adding 8 ml PBS, 4% paraformaldehyde (PFA) and overnight incu-
The cells are acquired (10,000 events) and analyzed with CellQuest Pro v3.3 on a Becton Dickinson FACSScalibur
flow cytometer. [0248] The OR and IN routes produce EboBunGP-specific neutralizing antibodies with the OR route producing the highest titres post-vaccination. The IM immunization produces detectable levels of neutralizing antibody. In comparison, 3/4 NHPs in the OR group demonstrate a 50% reduction in EboBun-GFP positive cells at a titre of 1:40. Similarly, the IN route results in a reduction of EboBun-GFP positive cells at the 1:40 dilution.

[0249] EboBunGP-specific effector cellular immune responses are determined using IL-2 and IFN-γ ELISPOT
assays as described by Qin, X. et al., PLoS ONE. 2009; 4(5):e5547 to determine the number of IL-2 and IFN-γ secreting lymphocytes. Prior to challenge on days 10 to 14 post-vaccination there is a detectable Ebo Bun immunogen-specific IFN-γ response in all immunized animals. The IM route is the most potent, inducing approximately 2-fold more IFN-γ secreting cells than OR (p<0.001) or IN (p=0.043) routes. A strong post-challenge secondary IFN-γ response is induced in all VSVAg/EboBun immunized animals with the IM route producing the most IFN-γ cells at day 6. By day 10 the OR group demonstrates a stronger response. The IFN-γ in the IN group rises steadily, peaking at day 26 post-challenge with 4.3 and 2 fold more EboBun specific IFN-γ secreting cells than the IM (p=0.003) and OR (p=0.075) group, respectively. All three routes produce strong EboBun-specific IFN-γ responses.

[0250] Post-vaccination, the IM group also has more EboBunGP-specific IL-2 secreting cells than either of the mucosally immunized groups. Post-challenge, the IM route continues to dominate early after challenge peaking on day 10. This difference shows a trend when compared to the IN group (p=0.067) and is significant when compared to the OR group (p<0.001). Additionally, the IN group has more IL-2 producing cells than the OR group (p=0.090) on day 10 post-challenge. By day 26 post-challenge all three routes continue to produce a EboBunGP-specific IL-2 response, however, the IN group response is strongest. At day 26 post-challenge the IN group has the most potent IFN-γ and IL-2 responses, as well as the highest IgA and IgG antibody titre, indicating this immunization route, followed by a EboBun challenge, results in the development of potent and sustained effector responses.

[0251] Absolute lymphocyte numbers for CD3+, CD4+, and CD8+ (CD3+, CD4+, CD8+) T cell populations are determined by flow cytometry. No decrease is observed in the lymphocyte populations for any of the VSVAg/EboBunGP vaccinated NHPs. In contrast, control animals who are not protected from EboBun show lymphocyte numbers decreased by 28-57%.

[0252] Macrophage numbers are slightly increased in control animals. However, the number of CD14+ cells is greater in the VSVAg/EboBunGP vaccinated groups with the IM route showing the most significant increases.

[0253] In order to determine the long term immune response after challenge, EboBunGP-specific CD4+ and CD8+ memory T-lymphocytes are examined for their ability to proliferate (CFSE) or produce IFN-γ in response to EboBunGP peptides at 6 months post-vaccination. EboBunGP-specific memory responses are observed as a result of vaccination followed by a ZEBOV challenge. These responses persist for at least 6 months. The memory populations in OR and IN inoculation routes demonstrate the greatest potential for proliferation and IFN-γ production post-challenge.

[0254] Any patents or publications mentioned in this specification are incorporated herein by reference to the same extent as if each individual publication is specifically and individually indicated to be incorporated by reference.

[0255] The compositions and methods described herein are presently representative of preferred embodiments, exemplary, and not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art. Such changes and other uses can be made without departing from the scope of the invention as set forth in the claims. All numerical ranges are inclusive of the whole integers and decimals between the endpoints, and inclusive of the endpoints.

REFERENCES


---

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 59
<210> SEQ ID NO 1
<211> LENGTH: 1894
<212> TYPE: DNA
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE: 
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Full viral sequence
<400> SEQUENCE: 1

cggcaccaca aaaaagtaa agagttttag aatctttatgg tgtgccagta acatacagga 60
agatttaaga ttttcctctt atgaaastg aatgtgat tctaatctcg acggatcgat 120
ccccacacc aacaactgaa atggctctgg agaagtattc tcgctcctgg caaacaag 180
agcgccgccc aagggccatt aggccacac ttcgctgagct gcgaagacag cagacattc 240
tcgccgaa agagcgcttg cccaaaacc ccaaaataa gcacaacgct ctcactcctga 300
ccccgattc ataaacata ttaagttgc aacagagata ctaaatcact aattgatca 360
agataaaaaa tagcggcaac actgccactc cttccctag aatgcaaga aatctttgt 420
agatctttgt gctacattct ctaaccaaga ctcctaaatg gacctctctg ctaactcag 480
tctgagatg ctaaaaacac ctagagttgc actaagactc actaagacac ccctggcccg 540
atgtccgctc cagcagggca ttcggagaca aagaacctt cccttttacc aataactcaca 600
cctgggaa gatgtcgaac tcatcataca gcattcagg gttggcctgc atttcaagga 660
tagtcgatg agtccttttg aattcctagct ttcgctctat gctatatcaag gggattataa 720
acaatccttg gaaagtaatg egtgtaataa ccctgaagat caattgattc cccttttggt 780
gagaaaaag gaaggtgtata gacgcctggaa gcagactctc cttccctgtct cagagttgaa 840
gaactcaag cagaaatctg tcgctaatcc cggagagaa cacaagcaag caaatgtggtg 900
acaatccttg ctaattgctgc gttgctttct caaagcagaattgg tgtgctgag aaaaagcttg 960
tctgagagag attcagagac aatcccttgtc gcctgcaaga cagagttgca tcaataacc 1020
gacacttctg caactcctgt ggacattatg gttctctctgc agaattcttc gcacaaatct 1080
cctgtaaaag ttcctctct ccatcagaaag aatgtgattg tgtcaggggt atgtcttcaaa 1140
tgatggtgtc attcagacatct gttcagttca aagccttttc ccgggatttg ggatgctttg 1200
aacaagttg tcatcataac ccaaaaataa aagcagcagga gttcctgctgc atccccctg 1260
gcgaacacgc caagtcctaaa aatggaattg ccttttaaac gctggccttc ggctcactag 1320
-continued

```
acaacatgga gaatagcccc cgggttgtcg tctgctgaaat ctatctgagg ttaataaatct
1380
tgagcatggt cttctctcct aactcttctgc aattgctttt ggagtttgc aatgctttag  ctgacatg  1440
gagcaacctg gctggagctca atgtaggaga gcaatacaca caaacctgcg aagoaaccac
1500
tgagggcggag aagacgtctg aagaaatgtgc aatcttctgt gcatatgatc aatcagcattc
1560
tgatatggct gaaagaagaa tcttaaaagta aacctcctcag aaaaagaatg agataagctt
1620
cagcagagcatg cacagcatgag tccacactgcg gaaagagaga tggcctcaat tgacgaagcg
1680
tattacctcc accttatcct ctaaaaccgag aagcgcgttt cttgagatca atgactattc
1740
ccttccaggg ccacataatg atacagagaa ctttggaaca aagcagtgacct cttcaacaga
1800
cctccaggat acaaactatt cggatgtact aatcagatc aaacagcggg ggttaaataa
1860
 ttacagctat tattcgaatt atgctgcaag tcgctcttga gctaatgctt tttttgacot
1920
tgagagagag atagatggctct ataacccgct ctaaaaccag cccagaaaaa atgataagcc
1980
agcacaacac aagctcgaaga atggagagcag ccagagatgg aacacagcac aacactctac
2040
ccagaggtga gcggccacacg ccacatcagca cctacacagac agcagcagagct
2100
cgacaaatctt gaaacccaga tttcaacacca gttgaggggg ttgtactctta tcagcgggga
2160
agcagacacg aagggccacaa agcagatgta atgtaagaag atctttctcctc ctgagactag
2220
cgagacaggt actgcaggtat ctacagtccag aagaacaaag cttgcactag ctcctctcgc
2280
tccgctctac gcagatcatt ccgagatgta tttgtcctcc tcagacaaca tcccgcacag
2340
gtccatcaaa acgaaacatg aggcactaggt cggcaacatg cttgagatgg aacacacagt
2400
tgagagagag tattactactc tttggaacac acagacaccc tttgtggcag tctctctata
2460
ccatctagct aagaaagagc cccatatattt cggcagactgt gatggggaag aagatacata
2520
tccagacctct cttgagatagt gatccmcaca cttgctctcag gagaagagag cctgatatac
2580
agacaaatga tttcaacacca gtagggctca gcagttttttac tcgctctgta gtaaatcag
2640
aataaatcct atggtgacttt tccgacttacg aatcagcgttc aagttgacatcc
2700
tcactacaag ctacccaatc actgtccgag aatgcacgaa cttccctcata aacacagctc
2760
acgacgaaac cccgctggag gtcacacgag caagacgagc acocctaata tgcctcttgt
2820
ccaccaacac cgcaacccag aactgtgcag tttccacacag aaotacaaac cccatggcag
2880
accaaccaac acggcaccac cccgataaca cccagccccc acacaaacca caaataacca ctgatgtgaa
2940
gtatgctcct cccccaggac gatcccttga cccacccaca cccctacact cccaaagtgg
3000
tcattagactt aataatagtt atgataaattc acacagtccgcc gcatgctctt ctcttctagc
3060
ttggtagctt ccagatctcc caaacactttc tataataact tcacattagct acocacttca
3120
agcacaacctg caagacctcc cccaccaaca cccagccagc cccacagctg tggggcgcgg
3180
aacacccggc cttgaccttt gcgcaaatga tggacgagcaaatggtgagct aacagatattc
3240
tcactaatct tgaactactta cttgtaaatgc tgcctcagct ccacactccaa atcaaaaacc
3300
ccaggttgtct aacacagctg gagcagagcc aacactgtcct aatgtgtaat gctgtattttg
3360
cagagatgtct gaaatgatac aactctctcta ccctggctgt acaaaaaacc acocctgcac
3420
tgactatact tgcagagacc aataactgc tggagatgag tggataaacc gcgtcctgaga
3480
tccacagactg tgcctctggta ctaaataagat tgcgtgtgaa gatgtgagcg ccacattagc
3540
ttcagtaaat gacgactgtggc cttcagactcg aacactgtgc agctactgaa gcatactggg
3600
```
-continued

cagaacatgg aagtctcca cccgggcct catattcog gaggatgcg atcaggacta 3660
aaattggaac aagggggtat atgtacccca aagagtcgca agagggcttc cgttaatcgg 3720
atagtaatgc cttcttacag gaagagaatt ttgggccacc agacataac gcacaaagct 3780
tgccacatg atcttgatag ccagctccag ggctgcgcac actattctat caactgtgca 3840
agattatgct caaggttagg aaggaacact cctcaactga tgtaatctat gcgaagattt 3900
agccagcctt tgctgaagag gacttccccg aagtgccctc gatccagata accaaacgga 3960
ctctatattt ccaagagctga gcacaccccc ctatctcatat tctgtcaagc ggtgtataac 4020
casaggccgtg tcacaaaggc ctccgccttg tctccacctc accaagagtt gataagaggtt 4080
gggtatgtcat ttgcaaggtc aagagccc gaaacactgg gactaaatat taaattaaggcc 4140
aatggtcgca ctctccatagt gctctctatt ggtctccg gcgaactagg gtagtccgag 4200
taaatataaatt tattgtactt cggctatc aactaatgac attctatgg ttctcgacag 4260
atagagagac gcagagcggct atcgagacc gatcttatgc ocaataccac taactgtttaa 4320
atatattata accaattgcct tctgttctta ttgatctttaa naaatcagacag ttaggtaagat 4380
taacaaaagct ccctttcactt ttcagagagct caatcttttt tatctctctct taactcttttt 4440
stgatgggtct ctcacctaccc actagagggc agatcttccct tctcgtccagg ccaaatatca 4500
tagaggttgtct tccgcacagtc agacatcgta gcctcttgtat caactgaagct ccagattgtc 4560
cyacatccct gacgccccag tgaatgtagc gtataaccct ccctcactcct ctcgccacca 4620
ttgctgtagta tataactagtc catcaaggiac ataccacccat acatggtcttc tcagccttta 4680
tacggtcgcc attgattagct gttaggtcgg gcggagagct acctattaagc caaatctctc 4740
tattgctctctctg ctccttattc atacagacta aacatatggc ttgattactc actacaggct 4800
caattatctg cggatccattt aacatcacttc aactggccaa aactctccat cggcttaattg 4860
gaatcactct aagtgctctct gggatccctc atacccgggt gcgtgctttta gcattagagaa 4920
atcagccgct ttgcttgcccct tgttgctgct ctcagccttc ctggcccgct gattattctct 4980
ctggatgcg ggccttaaag ctgtaactct aacatcttct gcgaagaccc ttggagggagct 5040
atcagccgct cgcggacttt gatccttgcct tcctctccat cccaaaactgt 5100
acctttacttt ttcgacgagg aagagccgga aagagggct tatctctgat tttatccttctc 5160
tcttctctattt ataagcactc atcagacccct caaacttcttg cegattgtcag 5220
cagccagga cattatatgct atctcgacag cgaaccttttt ctcacctccag 5280
agaaaaattc aacaaataaat ggtcaaacct tatctctactttt tctctaccac aagattattg 5340
geaatggatcc cttccatctct gcggacgccgc caatgttcgct caatgcagaac ttcacgcttc 5400
gecattctcttg tgcagtcttc ctgacactca gcgaagatta acatggaatg cggactggct 5460
cgcggccacc ggacctggtc ggcggctcag gcaccacacc caaaccctt ctggtccctt 5520
caaaaaacttt ttaaattctctt gactacgctc gacacctgctg gacacctgctg 5580
caacatctct acacacctgtc caattctgcc cgcacgcacct tagcagctctc 5640
tgctcagaaattg atccagcgca caaagtaacct caacactgca aggctaatct ggtgtattg 5700
aatctccgca cgcacccacac tacaactctt cccaaaattt ataatgttcgc gtcaacgccg 5760
acacttatgg tataaccttg ggcttccatt caagcgctctt cttactagg 5820
gtacccctct ctcagagata gtaatggctc ttccttataag aaaaaagaaaa 5880
cgaatgaagat taacgcgacc tagagegcgtg ccttcctcct ctaaatatttaa cttgtaata 5940
gagcaactca gtttggcaatt acctcctcttc cytagttgcac aaactcttcttcttgtaatt 6000
gtaataatac cacaagtcaat atggtaacat caggaatactc aacaaacgac ggtgaagct gcctct 6060
tcagaacccatat tttggggtaa taactcatttttcacaaagtt ttcctctctc 6120
catggcggt catccacaaac aacaactcttc cagtaaagttg tagaataaaataa gttggtgcc 6180
gggaataacac tttcctcaacag aacgagctga aacgtggcctggtcagtaca gaggaattg 6240
gagttggccgac agatgtacca accagcagagaa agagatggg aggacgacag ggttggctcc 6300
cacagtgtg gactgagcaa gttgatatgc cdfctacagac cgtgacacata cac 6360
tagagacag gttgtaggca tgcctacgcg aagccctgta ggggtaaga ggggctccc 6420
gctggcgtt gttgagacaga gttctcggaa cagggctgtg ccttgaaagtt taccggtttcc 6480
cacagaaaggg ggttttccct cttgactgtc atggcgtcata gacacattaac aacaactac tacatgatgas 6540
cacacttttc agaagatgtt ggtgccttctt tgactcctgc cggaaacaaa aagaacttttctt 6600
tcactagcgc acaactctctg ccaacgggcca atatgcaacgc agaccctctcc agaractacc 6660
acagctcc actataattc gttgcgtgac attttgggac caaatgtgaat aaccttctgtt 6720
ccataagggta ctcactctctg ttaaatgact cctggggcagg caaaggagtt cctggctgtgc 6780
tccactcctct ccactcctctc ggactaacaag attcaccaac caaatgtgtttctgtgtctgc 6840
tcacttcct cttggagactt ttaaatgact cctggggcagt ccagctgctca aagccctcctgatgtaa 6900
aaaaaacttt cccaaaaacc tttcacaagta agagcttgct ctgctaatggt taccoaagc 6960
ccagatcctc gggacgagaga aacagagccgg ccacctttctc accaggcctcc caacacccaa 7020
aacagagaga aacagagccgg ccacctttctc accaggcctcc caacacccaa 7080
agacctcccct gttcggagac cttgccccct ctggagcactac caggccacgc ggttcag 7140
acagctgcag ggacacgcgag aaaaaacc acaagaacta cgcacaaccag aacaacctcc 7200
aacaaccacaac gacagagcag gggcacttg gacagagcag gggcacttg gacagagcag 7260
cacaacatac ccgatgctgc cagccagact ttcctccatt cctccctctct ccagactgct 7320
tcctcctctcgc ccgtctccaca caagctgccg cctcccttcc ccagactgct 7380
cacaacaatg cacaacactc atgcacagcc cagcaactga cccaaaccct cttgcac oac 7440
cagactcatac gttgcaacac ctcactctctt acaccaaaaa ggggctgggg cttgacatg 7500
aggtctcaag agaaccggag gggaaacctc cttgagcca cagaaactt ctgaaacc 7560
cctcactcatt tggggtcag ctaaatatgtc gcctcagctcctt gttcttctgg 7620
ccaacttgggg cagacagagc gatatttaacgc ggaagggatatgca acagactcagc aaaaacttct 7680
aatttgggg ggcgggcttc attttttcagc ggggcatgctgtgtagtgc ggggagcagt 7740
tgcgtcagct gattgccgata ctttctctat atggactgcag aagagatgtac agtttttatct 7800
ccaaatgcttgg gccccccatgt gcactcttgc gagctgtgatagc cggggcatg 7860
ttcctcatac acaactattgc aacaatagta ctcaactctctgct ggtgactaatca tggtaaacc 7920
ttcctcatac acaactattgc aacaatagta ctcaactctctgct ggtgactaatca 7980
cagctgctgc gatcgggggc ttcctttcagc acggttagca ctggtagtgc cttgcaaat 8040
ttcacctata cttggcacc ctttgctaca ccctactg ggccttggag ctttggctaca 8100
ttcctctctct cttttctct tcaactgctac aagagatgtac gcctctctct 8160
aggggtgccg ggtgacttgg gtggggttg aggattgctc aatgtagacat cggccagaag gggatggtc 10500
tactgcttc gtttaaactc gctgacctcg ctctgctcgt gctgtgaca aggaatattg 10560
tttctcattt attcaaaact ccaaacttct acattgagtc tccctctcttg gcattagcag 10620
tgatcttggc aggctgttatt cagacaccag ttaatgcata acccctgttg 10680
cgggacccct gattgtttggt tcttttagaa aaaaacaaaa aatatttc 10740
tgcagaagcag gcgcgatcnaa gagcaagtga gctatcattt gtcggtctta 10800
atacatcgga attcatcagct caaattggaac ccacatagct tgtcgcacatt tgtgacact 10860
tgcagcattg caaattggaac actagaaactc atac ccacacaag aaccacacttg 10920
gttctctggt agaattaagc gagcctgtata cctctgctgca gaatcagaaag aaccacaccc 10980
cagtcgaatt cttctctcgt cagcataa cccgcttaa gttcatcttc aacattttagct 11040
tcacactgca ggcctgtgtt cttggaattta cagcctcctc cggcatagtt aacattttta 11100
ttttaatcagc caacctccaa ctgctctcct cttctatctc 11160
tacaagcagc aaccacttgc atttgaagaag tggggtcttt cttgcattt ctaattaaaca 11220
tatactata gattataaa aaccacataa cacaacctttc atacacatcG gagaatttc 11280
ggattgcctt aagataatac ctcgcatacta gagaagtttt cactcatttc tacaattttagct 11340
tatactata gttgatagct ctttctctca tatacctgct tggggtattc ttctttatatt gttttatatagc 11400
agtactaca tttacgagaa tctctatttc tgcagatgttc accctcctt atttgataatgct 11460
ttatactcagc taataagagc ttaatgcatt gaggctgaaa cttaaggataaa aaccacactg 11520
tctggtgatg aggcaacactg ccacaaattttc tctcttcatc ttttattagtt ttttattgct 11580
tacaacatc ccagactgca aaccattgct acccattttc ttcagatcgc ctttctttttt gtttattacatc 11640
caccagcttgt tggccgctgt actattttat cccgcatcagc tatactattc aacatttttg 11700
actacacaaaa atatatcata ggttcttttt tttatatagtt cttcattcagc caccacacta 11760
tgcagatcagc cttgagatatt cccaccttttc ttctttttatc caccacacta 11820
cgggagggc ttatgtttggt cggcagccat tttctgtgatt ctttacagc ggctagctgta 11880
ttatttttctt aggctgcagc tttttattat caccctactt ttgggggaggt 11940
tcagacagtt ggaaaattt ttgagcagc gattagcgttt gttatgattc ataatgattt 12000
tctgagcaca aggatttttt ggtcagtttt acggtatgtt ggtggttattt ggtgattttagt 12060
tcagagagta ccatgggattc aatcctgagc aaatacgatc ttattagatc tttcgggtta 12120
cggatttttt tttttcttttt cccctttttt ccttggttattt gggttattttt tattgggggatt 12180
tttcttttttttttt ggtggttattttt ggtggttattttt ggtggttattttt ggtggttattttt 12240
caccaacttc agoactagt ccgctgcaga ctattttttttt cccctttttttt ggtggttattttt 12300
tgggcttttctt ctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12360
tgcagtggtt cctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12420
gatctagttt cctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12480
gatctagttt cctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12540
ggctgaattc cacacaagtt aagagttttt tgggcttttt tgggtattttt aaccactattt cttttttttttttt 12600
tgggcttttctt ctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12660
tgggcttttctt ctgatctattt attgagttttt aaccactattt cttttttttttttt ggtggttattttt 12720
-continued

tattcagaaa gttaagaaaaa agtgaaccagt aataaaagca ttgcgtccccg tatacattc 12780
tgagacatat tgggtgttca agtacagcat tggcaacact tatitgata gcaccaggtc 12840
atggtatagt gtaatctcag ataaraacat aacaccoggles ttacacttct acattaagag 12900
gaacacattt cggccactcgct tattgctaaag actacctattgg tgggaatttct atcacccttga 12960
ttcattctcc tttatttttca ccagacttat tattgacttg agtatatcctttaaag gaatgc 13020
cgactcagca gtggaaaaaaa catgttggga tgcagtttctt gagcctaatgt tttctgtgata 13080
tagttctccaa aacaagtactt caetaaagag gtctctgtta aagttacttgg aaccaagaaa 13140
ttttcctgatt gatgctgttc ttaatattgac ccagccctctg gattataac acataataacata 13200
ccggtaatattt ttttttttat ctaagaaaaa aagattaattt gtagagacag gtttttgtat 13260
gctaaacctct ccctacagtt taatgtgtaa aacgctattgtc cagatggattatgacgactgg 13320
agctaaagcc tttccatagta acatgatgcttg tataaaagag cgtgacgaca aggaaagcct 13380
cttgaaccccg ggtgctggcc aaccaccaag tgcacattcctt ggtggaattg ccactgttag 13440
agcagcagcttttcttgctag tccaagaaaa cacatcactgt gcaatgtgtcttac 13500
agctcatttat agtacctact gtaatgtgta ttatgctttaa aaaaatattt tcaatctggt 13560
gcattataac atacccaatg gttattacaa ttagaatgtag attattataa cccctcactgg 13620
agttcagcta gaaatactcggg aagatcecore gcaagccacct agcctttactgt ggtgctcct 13680
tgggggactgttg gagggacctcc aaccaagaaa ctgggacgcgct attccatggt cacaaaacctcc 13740
attatgtgag atcaagagctttt ctaacattgt gattctcctgct gtaatggggt ctaaactaatg 13800
catcacaagttttctcctgt tttctctgtaa ccagatctcttc atgacgcaag agatagtcg 13860
cgaggacaatt ggtgctggcag tagacgcaag tttaccaagga gtoacaggtt ctgtgctcat 13920
tctctacaaaa cccagatgaga cttttgtgca ttcaggctttt attatattctgtataagagca 13980
statatataat ggtgcttaat tgcacaacca actaaagactg acataccagga ttgttccacc 14040
gtcaagagcag atcttgatgag accttatggcg aacctgataag gccgattttg 14100
aggacttataa ttcggaacta gacatgtataa cccctggcctc gttgctggcgc cattctcac 14160
atatctttttcgcttttgcaaa cttcattttctg ttcacaaaaa cacaccagact 14220
aggtcaactta tcatactgaa aacccgctgct ttcctgcaact aatctctttg ctttacgcggt 14280
acccagattctacagagtt actacaggcttatttttaa aacaccagag aatatttttat cggccact 14340
tgggaccccg ggacccctccgg gctttatcacta actagctgaat tattggccaa acatccacat 14400
ggaagctutta tttcttctaggt tatttggccag aaccacccggg aactcttagtgc caattgtact 14460
tgcaacctac caacacgctgtaagcttctc ttgggtaacaa gaccaaaatct cttttttaegt 14520
tcattaagagctgagcataaactcatttactg cttattatagta aacgtatatatttatggagtt 14580
tcactctcat gcaggtttttaga aagagtagagtg gtagctaaaat tgcctactcctt ttaaacaac 14640
ttaattaggt ggtgtgctggt atctatataa cctcgcgttact cgcctttgctagagctggaa 14700
gatctctgct ttgttataagag ggaccaagct ttggctatcctt ctaacactgct ttaaaaattctt 14760
tgggaagcttactttagag ataggttgag gtaaatacctc atgcaagagct ggtttttgtgtgct 14820
gtttagctcat ctagaccaagct gttgagcagt tttctagggt aaattttaaa aagtttttcttg 14880	taggcttggt ggccagcaaa tttacgttgaata acacttccgag ccacaacata tagggagag 14940
aacgtcttact ggtgcaacacct ttcctctgcct gttgagaaaaa tttataagctgcttggctca 15000
atcgtacggaa caagtcccta aatgtgacaa atctagaat ccaaaaggag agccatattg 15060
gtcaattggca attaaagaaac aagtttgtag tgcagggcag matcgtcac ggttaatttg 15120
gacacattgg gaagctgtac ttccatcagg gcttcgaaca gaggcaaga tggggcagcc 15180
agcactcaac ctcagttgcct cctcgtgctc ctctcggcga gcataagagt tgtactcctag 15240
actacactgg gttaccccaag tgggtgcacaa tagttatgg ttagttaaac ctttttgaga 15300
gcagagagca accttgaagt tgtgaggagat ctttaaataat agcgccttctc attttccag 15360
gacacgtgca ctcggtacta attaccaata cagccctctct tctttcattg caaatagaaa 15420
gagttaatcgc gcaagcgagat tgtgtggtgct gcaaaatcaact cttccgggatt ttcctttgag 15480
ggggcaactca gcagcgccgac gcacactaatt ctctttaaacatgtcaacttt ctctacaogt 15540
ctctattgat ttaacagtctc gagaaaccga aacaactcctc attacacata attgtgcgcca 15600	
ttctctcttt tttcagatttgg gcacccggaga aagcgcctcgt ctaatactta cctacaagct 15660
tagcttttccc ttggagttcga caaaggacg agaagatag ttaattattgt atacaactcc 15720
gtaaaaaggt ggcctttact gcaccacttcc ctttttcttc aagggcgaag 15780
gtccacatc atagagggag atttgattag attcttacct ctatcctggt gggaacctgc 15840
gaaacaatc attcaagctta tattctcagag caagacttacatct tctccacaag accccattag 15900
cagttgagaa agccgcatcatt caacacatctt cttttctcaat cttctcatttg agggtctctt 15960
tcattgggtcc gcagagagat tgtggtatagct aagcgcctttctt ccctttttta aagggcgaag 16020
gotagaacctcg tgtactttctg ttaacaactta aacaactcctag atacataatg ttgcaactcg 16080
tcgtctgtagg atctcagcatt ccaacctttca aacagttagtgt gtgatatcaac gacacatagt 16140	
tattgcctt catttttctcc ctataccaggg gggtaaggca ggtgatcag ggtcctctggga 16200
tgctcagcaga ctcattgctct cgtgcctcttt ctctaattct tcaaaaatagt 16260
gatcgtbgaa taacagagac ctaactctct tgtgtgatatt cctccctcatttg aagggarcas 16320
tccagattca attaagctt tttcactcgt gattatgccc ttaagtgcaga agtaactccc 16380
tcacaacga acacacactc aacagagac caataataca cagttgctct ttatactcattg 16440
tcctactctgc aagcgcctct cccttattt ccctcctgcac tcactctgcct atggtggagg 16500
cggacacaa gagcggccaa aaaaatggat gacggaagaa cagacagtta aacocatacc 16560
atatggacta tttcattctc ttaataattgc ccaaaaccca ccaagcactc ccaaatccta 16620
gtcgaagaact caaggtctca ggcgcttttt tgcggaacct gatagatgta aagaagagaa 16680
attgocacac gttcgtcgac gacggagcacc atcagagacc tataatcacc ggccctcagc 16740
cgaaattcat ctcgtaaaac caccatcataa tgcgctaataaagagatcaag tgtactcctag 16800
gggtggtgat tcaagttcag aaaaattgct tgaagtctcct aacaagccctt aatatctg 16860
tttcacttg ttcacgagcc actacagaaact cttctctgct aaaaaacttgctagtatataac 16920	
tgaaataaacc aacatcattt gacaattatca agaagcttccag atagctttgac 16980
tttaaagctct ctcagctggt cttaaatgtt gaggctttctc gggaattcga 17040
tagtttaaacc acctgggtgta cctcagctggt aggaaacgg gtcagttctcctt tagttactctg 17100
acaacaatatt aaggtcagaa caatcttttt taacaactctta atacagagct atacatcga 17160
ggagaaaata gttttcggga ccaaccccac tcatgtctgc cttcctgttgcagctgccaca 17220	
ttattgagtc caataataaac ttaaatcttct ccaattgtctg agcggacgtta ctgatatcgc 17280
-continued

taacctgca tggttcactg accagaatac tagaatcccc accacaagtg agattatgc a 17340
tatgatgtc gaaacgcag aaaaattaa ttgctaaaa ttatatgagc cttatcagca 17400
attaatttt tcacaccatgt ataoaagggt gotaaagatt gttattataa aggtttttttt 17460
agatgtatat gacagcttcc tgtgctttaa tgcacccctt gccccttttat tgcgggcgg 17520
catttaatt aacatatta ottogagtc aatgtcaggc gatgtctactc atgttcttc 17580
aaattcctc tcaagcttcc gacgaaggcc tctagaaggct cagtcctactgtc agatgcaag 17640
catcacaaca gcgcgtacag tcaccagttca aagaggttca tcatgctttta gcacattgtg 17700
gcaatgtct gatattatt tgtacacttg attgtattag tgtgattttcc cttctattgga 17760
aaagttttct taccactogat atacactagt gtatgtoacc gcgggtttcc tgcctctgat 17820
cctttaaggct ttaaactcat tgcagcagaa gattgagaa ttatgtggtgacctatacta 17880
gcaagacaag atgtcaaac caaactataa otcctcataa acgcaacagg acgcgttac 17940
aaataaggct aatgtcaact cttatttaatt ccctctggttg caagactaca agatattctt 18000
tctttggcag gcagacgtact gaaaccttcc tgcctattac aatgtttgca aatgatatatta 18060
cataaatgg gcgtgctcact gttgagctcg atttataatt cccatactttt caataaaccg 18120
tgctcaacgc tccaagaaca aatattgc gataaatcc gcgggtcttag gattgatatcc 18180
tatctgtat gacagcttcc acacaccaat tagactatcc aacacccttaa 18240
tggataatt ctcctcattcg acagagggagc ccattcttatt aacacattta 18300
caacagctg cagctgcttg ttttacgac aacacaggtg gcgggtcaca tcaggttcac 18360
ttgtgatgt ttccttcacta tgcagctacc ctcataatgtt aatggttatc cttcttatt 18420
cctatgacc gctacactctg ccaataactct atcaataag aaaaattact gttggttctg 18480
cattttatg acatcagca gcgtatatat atatatttact cttacctaaa tgtgttagtt 18540
tagtaattgg aaggttataatt taggtattag gtttactata taataaatagt cactatagg 18600
tatccacaag aagtgtagtat tttgtcacta aataatattc acagaggttt gttttattaa 18660
atataaagct tattgtgtg gttggtgttg gtaataaatct gttatctttca 18720
cctataactatat attatatag tagatcagta tgcgttctgg cttatcacaac cttagaaaaa 18780
tggtgagct gtttttacca tgcacacatt gattgagggt getttcattac 18840
aaatgctga attaatcttcga atgtaggtgga gacacacttt ggttatagttg tcgacacaa 18900
aatatcctttt caaattgtgtttttttctttctgttttcca 18940

<210> SEQ ID NO: 2
<211> LENGTH: 2210
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<222> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus L viral protein

<400> SEQUENCE: 2

Met Ala Thr Gln His Thr Gln Tyr Pro Asp Ala Arg Leu Ser Ser Pro 1 5 10 15
Ile Val Leu Asp Gln Cys Asp Leu Val Thr Arg Ala Cys Gly Leu Tyr 20 25 30
Ser Ser Tyr Ser Leu Asn Pro Gln Leu Lys Asn Cys Arg Leu Pro Lys 35 40 45
His Ile Tyr Arg Leu Lys Phe Asp Ala Thr Val Thr Lys Phe Leu Ser
50 55 60
Asp Val Pro Ile Val Thr Leu Pro Ile Asp Tyr Leu Thr Pro Leu Leu
65 70 75 80
Leu Arg Thr Leu Ser Gly Glu Gly Leu Cys Pro Val Glu Pro Lys Cys
95 100 105 110
Ser Gln Phe Leu Asp Glu Ile Val Ser Tyr Val Leu Gln Asp Ala Arg
120 125 130 135
Phe Leu Arg Tyr Phe Arg His Val Gly Val His Asp Asp Asn Val
140 145 150 155 160
Gly Lys Asn Phe Glu Pro Lys Ile Lys Ala Leu Ile Tyr Asp Asn Glu
170 175 180 185 190
Phe Leu Gln Gln Leu Phe Tyr Trp Tyr Asp Leu Ala Ile Leu Thr Arg
195 200 205 210 215
Arg Gly Arg Leu Asn Arg Gly Asn Arg Ser Thr Trp Phe Ala Asn
220 225 230 235 240
Arg Asp Leu Ile Asp Ile Leu Gly Tyr Gly Asp Tyr Ile Phe Trp Lys
245 250 255 260 265 270
Ile Pro Leu Ser Leu Leu Ser Leu Thr Glu Gly Ile Pro His Ala
275 280 285 290 295
Ala Lys Asp Trp Tyr His Ala Ser Ile Phe Lys Glu Ala Val Gln Gly
300 305 310 315 320
His Thr His Ile Val Ser Val Ser Thr Ala Asp Val Leu Ile Met Cys
325 330 335
Lys Asp Ile Ile Thr Cys Arg Phe Asn Thr Thr Leu Ile Ala Ala Leu
340 345 350 355 360 365
 Ala Asn Leu Glu Asp Ser Ile Cys Ser Asp Tyr Pro Gln Pro Glu Thr
370 375 380 385 390 395 400
Ile Ser Asn Leu Tyr Lys Ala Gly Asp Tyr Leu Ile Ser Ile Leu Gly
395 400 405 410 415 420 425 430
Ser Glu Gly Tyr Lys Val Ile Lys Phe Leu Glu Pro Leu Cys Leu Ala
435 440 445
Pro Pro Leu Pro Met Ile Lys Asp Leu Leu Trp Glu Phe Tyr His Leu
450 455
Aasp His Pro Pro Leu Phe Ser Thr Lys Ile Ile Ser Asp Leu Ser Ile
465 470 475 480
Phe Ile Lys Asp Arg Ala Thr Ala Val Glu Lys Thr Cys Trp Asp Ala
485 490 495
Val Phe Glu Pro Asn Val Leu Gly Tyr Ser Pro Pro Asn Lys Phe Ser
500 505 510
Thr Lys Arg Val Pro Glu Gin Phe Leu Glu Gin Glu Asn Phe Ser Ile
515 520 525
Aasp Ser Val Leu Thr Tyr Ala Gin Arg Leu Asp Tyr Leu Leu Pro Gin
530 535 540
Tyr Arg Asn Phe Ser Phe Ser Leu Lys Glu Lys Glu Leu Asn Val Gly
545 550 555 560
Arg Ala Phe Gly Lys Leu Pro Tyr Pro Thr Arg Asn Val Gin Thr Leu
565 570 575
Cys Glu Ala Leu Leu Ala Asp Gly Leu Ala Lys Ala Phe Pro Ser Asn
580 585 590
Met Met Val Val Thr Glu Arg Glu Gin Lys Glu Ser Leu Leu His Gin
595 600 605
Ala Ser Trp His His Thr Ser Asp Asp Phe Gly Gin Asn Ala Thr Val
610 615 620
Arg Gly Ser Ser Phe Val Thr Asp Leu Glu Lys Tyr Asn Leu Ala Phe
625 630 635 640
Arg Tyr Glu Phe Thr Ala Pro Phe Ile Glu Tyr Cys Asn Arg Cys Tyr
645 650 655
Gly Val Lys Asn Leu Phe Asn Trp Met His Tyr Thr Ile Pro Gin Cys
660 665 670
Tyr Ile His Val Ser Asp Tyr Tyr Asn Pro Pro His Gly Val Ser Leu
675 680 685
Glu Asn Arg Glu Asp Pro Pro Glu Gly Pro Ser Ser Tyr Arg Gly His
690 695 700
Leu Gly Ile Glu Gly Leu Gin Gin Lys Leu Trp Thr Ser Ile Ser
705 710 715 720
Cys Ala Gin Ile Ser Leu Val Glu Ile Lys Thr Gly Phe Lys Leu Arg
725 730
Ser Ala Val Met Gly Asp Asn Gin Cyc Ile Thr Val Leu Ser Val Phe
740 745 750
Pro Leu Glu Thr Asp Ser Asn Glu Gin Gin His Ser Ser Gly Asp Asn
755 760 765
Ala Ala Arg Val Ala Ala Ser Leu Ala Lys Val Thr Ser Ala Cys Gly
770 775 780
Ile Phe Leu Lys Pro Asp Glu Thr Phe Val His Ser Gly Phe Ile Tyr
785 790 795 800
Phe Gly Lys Lys Gin Tyr Leu Asn Gly Val Gin Leu Pro Gin Ser Leu
805 810 815
Lys Thr Ala Thr Arg Ile Ala Pro Leu Ser Asp Ala Ile Phe Asp Asp
820 825 830
Leu Gin Gly Thr Leu Ala Ser Ile Gly Thr Ala Phe Glu Arg Ser Ile
835 840 845
Ser Glu Thr Arg His Val Tyr Pro Cys Arg Val Val Ala Ala Phe His
Thr Phe Phe Ser Val Arg Ile Leu Gln Tyr His His Leu Gly Phe Asn
Lys Gly Thr Asp Leu Gly Gln Leu Ser Leu Ser Lys Pro Leu Asp Phe
Gly Thr Ile Thr Leu Ala Leu Ala Val Pro Gin Val Leu Gly Gln Leu
Ser Phe Leu Asn Pro Glu Lys Cys Phe Tyr Arg Asn Leu Gly Asp Pro
Val Thr Ser Gly Leu Phe Gin Leu Arg Thr Tyr Leu Gin Met Ile Asn
Met Asp Asp Leu Phe Leu Pro Leu Ile Ala Lys Asn Pro Gly Asn Cys
Ser Ala Ile Asp Phe Val Leu Asn Pro Ser Gly Leu Asn Val Pro Gly
Ser Gin Asp Leu Thr Ser Phe Leu Arg Gin Ile Val Arg Arg Thr Ile
Thr Leu Ser Ala Lys Asn Lys Leu Ile Asn Thr Leu Phe His Ser Ser
Ala Asp Leu Glu Asp Glu Met Val Cys Lys Trp Leu Leu Ser Ser
Thr Pro Val Met Ser Arg Phe Ala Ala Asp Ile Phe Ser Arg Thr
Pro Ser Gly Lys Arg Leu Gin Ile Leu Gly Tyr Leu Glu Gly Thr
Arg Thr Leu Leu Ala Ser Lys Val Ile Asn Asn Ala Glu Thr
Pro Ile Leu Asp Arg Leu Arg Lys Ile Thr Leu Gin Arg Trp Ser
Leu Trp Phe Ser Tyr Leu Asp His Cys Asp Gin Val Leu Ala Asp
Ala Leu Ile Lys Val Ser Cys Thr Val Asp Leu Ala Gin Ile Leu
Arg Glu Tyr Thr Trp Ala His Ile Leu Glu Gly Arg Gin Leu Ile
Gly Ala Thr Leu Pro Cys Met Leu Glu Gin Phe Asn Val Phe Trp
Leu Lys Ser Tyr Glu Gin Cys Pro Lys Cys Ala Lys Ser Arg Asn
Pro Lys Gly Glu Pro Phe Val Ser Ile Ala Ile Lys Lys Gin Val
Val Ser Ala Thr Pro Asn Gin Ser Arg Leu Asn Thr Ile Gly
Asp Gly Val Pro Tyr Ile Gly Ser Arg Thr Glu Asp Lys Ile Gly
Gln Pro Ala Ile Lys Pro Lys Cys Pro Ser Ala Ala Leu Arg Glu
Ala Ile Glu Leu Thr Ser Arg Leu Thr Trp Val Thr Gin Gly Gly
Ala Asn Ser Asp Leu Leu Val Lys Pro Phe Val Glu Ala Arg Val
<table>
<thead>
<tr>
<th>Leu</th>
<th>Val</th>
<th>Tyr</th>
<th>Met</th>
<th>Cys</th>
<th>Lys</th>
<th>Ser</th>
<th>Thr</th>
<th>Ala</th>
<th>Ser</th>
<th>Asn</th>
<th>Phe</th>
<th>Phe</th>
<th>His</th>
<th>Ala</th>
<th>1625</th>
<th>1630</th>
<th>1635</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
<td>Leu</td>
<td>Ala</td>
<td>Tyr</td>
<td>Trp</td>
<td>Arg</td>
<td>Ser</td>
<td>Arg</td>
<td>His</td>
<td>Lys</td>
<td>Gly</td>
<td>Arg</td>
<td>Pro</td>
<td>Lys</td>
<td>Asn</td>
<td>1640</td>
<td>1645</td>
<td>1650</td>
</tr>
<tr>
<td>Arg</td>
<td>Ser</td>
<td>Thr</td>
<td>Glu</td>
<td>Glu</td>
<td>Gln</td>
<td>Thr</td>
<td>Val</td>
<td>Lys</td>
<td>Pro</td>
<td>Ile</td>
<td>Pro</td>
<td>Tyr</td>
<td>Asp</td>
<td>Asn</td>
<td>1655</td>
<td>1660</td>
<td>1665</td>
</tr>
<tr>
<td>Phe</td>
<td>His</td>
<td>Ser</td>
<td>Val</td>
<td>Lys</td>
<td>Cys</td>
<td>Ala</td>
<td>Ser</td>
<td>Asn</td>
<td>Pro</td>
<td>Pro</td>
<td>Ser</td>
<td>Ile</td>
<td>Pro</td>
<td>Lys</td>
<td>1670</td>
<td>1675</td>
<td>1680</td>
</tr>
<tr>
<td>Ser</td>
<td>Lys</td>
<td>Ser</td>
<td>Gly</td>
<td>Thr</td>
<td>Glu</td>
<td>Gln</td>
<td>Gly</td>
<td>Ser</td>
<td>Ser</td>
<td>Ala</td>
<td>Phe</td>
<td>Phe</td>
<td>Glu</td>
<td>Lys</td>
<td>Leu</td>
<td>1685</td>
<td>1690</td>
</tr>
<tr>
<td>Glu</td>
<td>Tyr</td>
<td>Asp</td>
<td>Lys</td>
<td>Glu</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Pro</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
<td>Thr</td>
<td>Pro</td>
<td>Ala</td>
<td>1700</td>
<td>1705</td>
<td>1710</td>
</tr>
<tr>
<td>Glu</td>
<td>Gln</td>
<td>Ser</td>
<td>Lys</td>
<td>Thr</td>
<td>Tyr</td>
<td>Ile</td>
<td>Lys</td>
<td>Ala</td>
<td>Ser</td>
<td>Ser</td>
<td>Arg</td>
<td>Ile</td>
<td>Tyr</td>
<td>1715</td>
<td>1720</td>
<td>1725</td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Gly</td>
<td>Lys</td>
<td>Thr</td>
<td>Pro</td>
<td>Ser</td>
<td>Asn</td>
<td>Ala</td>
<td>Ala</td>
<td>Lys</td>
<td>Asp</td>
<td>Ser</td>
<td>Thr</td>
<td>Thr</td>
<td>1730</td>
<td>1735</td>
<td>1740</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Lys</td>
<td>Gly</td>
<td>Cys</td>
<td>Asp</td>
<td>Ser</td>
<td>Lys</td>
<td>Glu</td>
<td>Asn</td>
<td>Ala</td>
<td>Val</td>
<td>Gln</td>
<td>Ala</td>
<td>Ser</td>
<td>1745</td>
<td>1750</td>
<td>1755</td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Arg</td>
<td>Ile</td>
<td>Val</td>
<td>Leu</td>
<td>Pro</td>
<td>Phe</td>
<td>Phe</td>
<td>Thr</td>
<td>Leu</td>
<td>Ser</td>
<td>Gln</td>
<td>Asn</td>
<td>Asp</td>
<td>Tyr</td>
<td>1760</td>
<td>1765</td>
<td>1770</td>
</tr>
<tr>
<td>Arg</td>
<td>Thr</td>
<td>Pro</td>
<td>Ser</td>
<td>Ala</td>
<td>Lys</td>
<td>Ser</td>
<td>Glu</td>
<td>Tyr</td>
<td>Ile</td>
<td>Thr</td>
<td>Glu</td>
<td>Ile</td>
<td>Thr</td>
<td>1775</td>
<td>1780</td>
<td>1785</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Leu</td>
<td>Ile</td>
<td>Arg</td>
<td>Gln</td>
<td>Leu</td>
<td>Lys</td>
<td>Ala</td>
<td>Ile</td>
<td>Pro</td>
<td>Asp</td>
<td>Thr</td>
<td>Val</td>
<td>Tyr</td>
<td>1790</td>
<td>1795</td>
<td>1800</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Arg</td>
<td>Phe</td>
<td>Thr</td>
<td>Gly</td>
<td>Val</td>
<td>Val</td>
<td>Ser</td>
<td>Ser</td>
<td>Met</td>
<td>His</td>
<td>Tyr</td>
<td>Lys</td>
<td>Leu</td>
<td>Asp</td>
<td>1805</td>
<td>1810</td>
<td>1815</td>
</tr>
<tr>
<td>Glu</td>
<td>Val</td>
<td>Leu</td>
<td>Trp</td>
<td>Glu</td>
<td>Phe</td>
<td>Asp</td>
<td>Ser</td>
<td>Phe</td>
<td>Lys</td>
<td>Thr</td>
<td>Ala</td>
<td>Val</td>
<td>Thr</td>
<td>Leu</td>
<td>1820</td>
<td>1825</td>
<td>1830</td>
</tr>
<tr>
<td>Ala</td>
<td>Glu</td>
<td>Gly</td>
<td>Glu</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
<td>Ala</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Gln</td>
<td>Lys</td>
<td>Tyr</td>
<td>1835</td>
<td>1840</td>
<td>1845</td>
</tr>
<tr>
<td>Lys</td>
<td>Val</td>
<td>Arg</td>
<td>Thr</td>
<td>Ile</td>
<td>Phe</td>
<td>Phe</td>
<td>Asn</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
<td>His</td>
<td>Ser</td>
<td>1850</td>
<td>1855</td>
<td>1860</td>
</tr>
<tr>
<td>Ile</td>
<td>Glu</td>
<td>Ala</td>
<td>Glu</td>
<td>Ile</td>
<td>Val</td>
<td>Ser</td>
<td>Gly</td>
<td>Thr</td>
<td>Thr</td>
<td>Thr</td>
<td>Pro</td>
<td>Arg</td>
<td>Met</td>
<td>Leu</td>
<td>1865</td>
<td>1870</td>
<td>1875</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Val</td>
<td>Met</td>
<td>Ala</td>
<td>Lys</td>
<td>Leu</td>
<td>His</td>
<td>Asp</td>
<td>Asp</td>
<td>Gln</td>
<td>Ile</td>
<td>Asn</td>
<td>Val</td>
<td>Ile</td>
<td>1880</td>
<td>1885</td>
<td>1890</td>
</tr>
<tr>
<td>Leu</td>
<td>Aen</td>
<td>Aen</td>
<td>Ser</td>
<td>Ala</td>
<td>Ser</td>
<td>Gln</td>
<td>Val</td>
<td>Thr</td>
<td>Asp</td>
<td>Ile</td>
<td>Thr</td>
<td>Aen</td>
<td>Pro</td>
<td>Ala</td>
<td>1895</td>
<td>1900</td>
<td>1905</td>
</tr>
<tr>
<td>Trp</td>
<td>Phe</td>
<td>Thr</td>
<td>Asp</td>
<td>Gln</td>
<td>Ser</td>
<td>Arg</td>
<td>Ile</td>
<td>Pro</td>
<td>Thr</td>
<td>Gln</td>
<td>Val</td>
<td>Glu</td>
<td>Ile</td>
<td>1910</td>
<td>1915</td>
<td>1920</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Thr</td>
<td>Met</td>
<td>Asp</td>
<td>Ala</td>
<td>Glu</td>
<td>Thr</td>
<td>Thr</td>
<td>Glu</td>
<td>Asn</td>
<td>Ile</td>
<td>Asn</td>
<td>Arg</td>
<td>Ser</td>
<td>Lys</td>
<td>1925</td>
<td>1930</td>
<td>1935</td>
</tr>
<tr>
<td>Leu</td>
<td>Tyr</td>
<td>Glu</td>
<td>Ala</td>
<td>Ile</td>
<td>Gln</td>
<td>Leu</td>
<td>Ile</td>
<td>Val</td>
<td>Ser</td>
<td>His</td>
<td>Ile</td>
<td>Asp</td>
<td>Thr</td>
<td>1940</td>
<td>1945</td>
<td>1950</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Val</td>
<td>Leu</td>
<td>Lys</td>
<td>Ile</td>
<td>Val</td>
<td>Ile</td>
<td>Lys</td>
<td>Val</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
<td>Asp</td>
<td>Ile</td>
<td>1955</td>
<td>1960</td>
<td>1965</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Tyr</td>
<td>Leu</td>
<td>Lys</td>
<td>Pro</td>
<td>Ile</td>
<td>Thr</td>
<td>Ser</td>
<td>Ser</td>
<td>Pro</td>
<td>Lys</td>
<td>Ser</td>
<td>Ser</td>
<td>1985</td>
<td>1990</td>
<td>1995</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Trp</td>
<td>Tyr</td>
<td>Leu</td>
<td>Cys</td>
<td>Leu</td>
<td>Ser</td>
<td>Asn</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td>Ser</td>
<td>Arg</td>
<td>1995</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 3
<211> LENGTH: 739
<212> TYPE: PRO
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus NP viral protein
<400> SEQUENCE:
Met Asp Pro Arg Pro Ile Arg Thr Trp Met Met His Asn Thr Ser Glu 1 5 10 15
Val Glu Ala Asp Tyr His Lys Ile Leu Thr Ala Gly Leu Ser Val Gln 20 25 30
Gln Gly Ile Val Arg Gln Arg Ile Pro Val Tyr Gln Ile Ser Asn 35 40 45
Leu Glu Glu Val Cys Glu Ile Glu Glu Ala Phe Glu Ala Gly Val 50 55 60
Asp Phe Glu Asp Ser Ala Asp Ser Phe Leu Leu Met Leu Cys Leu His 65 70 75 80
His Ala Tyr Gln Gly Asp Tyr Lys Gln Phe Leu Glu Ser Asn Ala Val 85 90 95
Lys Tyr Leu Glu Gly His Gly Phe Arg Phe Glu Met Lys Lys Glu 100 105 110
<table>
<thead>
<tr>
<th>Gly</th>
<th>Val</th>
<th>Lys</th>
<th>Arg</th>
<th>Leu</th>
<th>Glu</th>
<th>Glu</th>
<th>Leu</th>
<th>Leu</th>
<th>Pro</th>
<th>Ala</th>
<th>Ala</th>
<th>Ser</th>
<th>Ser</th>
<th>Gly</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Ile</td>
<td>Lys</td>
<td>Arg</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Met</td>
<td>Pro</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Thr</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asn</td>
<td>Ala</td>
<td>Gly</td>
<td>Gln</td>
<td>Phe</td>
<td>Leu</td>
<td>Ser</td>
<td>Phe</td>
<td>Ala</td>
<td>Ser</td>
<td>Leu</td>
<td>Phe</td>
<td>Leu</td>
<td>Pro</td>
<td>Lys</td>
</tr>
<tr>
<td>145</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Val</td>
<td>Val</td>
<td>Gly</td>
<td>Glu</td>
<td>Lys</td>
<td>Ala</td>
<td>Cys</td>
<td>Leu</td>
<td>Glu</td>
<td>Lys</td>
<td>Val</td>
<td>Gln</td>
<td>Arg</td>
<td>Gln</td>
<td>Ile</td>
</tr>
<tr>
<td>165</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Val</td>
<td>His</td>
<td>Ala</td>
<td>Glu</td>
<td>Gln</td>
<td>Gly</td>
<td>Leu</td>
<td>Ile</td>
<td>Gln</td>
<td>Tyr</td>
<td>Pro</td>
<td>Thr</td>
<td>Ser</td>
<td>Trp</td>
<td>Gln</td>
</tr>
<tr>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Gly</td>
<td>His</td>
<td>Met</td>
<td>Met</td>
<td>Val</td>
<td>Ile</td>
<td>Phe</td>
<td>Arg</td>
<td>Leu</td>
<td>Met</td>
<td>Arg</td>
<td>Thr</td>
<td>Asn</td>
<td>Phe</td>
</tr>
<tr>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ile</td>
<td>Lys</td>
<td>Phe</td>
<td>Leu</td>
<td>Leu</td>
<td>Ile</td>
<td>His</td>
<td>Gln</td>
<td>Gly</td>
<td>Met</td>
<td>His</td>
<td>Met</td>
<td>Val</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Asp</td>
<td>Ala</td>
<td>Asn</td>
<td>Asp</td>
<td>Ala</td>
<td>Val</td>
<td>Ala</td>
<td>Asn</td>
<td>Ser</td>
<td>Val</td>
<td>Ala</td>
<td>Gln</td>
<td>Ala</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td>225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Ser</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td>Ile</td>
<td>Val</td>
<td>Lys</td>
<td>Thr</td>
<td>Val</td>
<td>Leu</td>
<td>Asp</td>
<td>His</td>
<td>Ile</td>
<td>Leu</td>
<td>Gln</td>
</tr>
<tr>
<td>245</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Thr</td>
<td>Glu</td>
<td>His</td>
<td>Gly</td>
<td>Val</td>
<td>Arg</td>
<td>Leu</td>
<td>His</td>
<td>Pro</td>
<td>Leu</td>
<td>Ala</td>
<td>Arg</td>
<td>Thr</td>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Lys</td>
<td>Asn</td>
<td>Glu</td>
<td>Val</td>
<td>Ser</td>
<td>Ser</td>
<td>Phe</td>
<td>Lys</td>
<td>Ala</td>
<td>Ala</td>
<td>Leu</td>
<td>Ala</td>
<td>Ser</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td>275</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>His</td>
<td>Gly</td>
<td>Glu</td>
<td>Tyr</td>
<td>Ala</td>
<td>Pro</td>
<td>Phe</td>
<td>Arg</td>
<td>Leu</td>
<td>Asn</td>
<td>Leu</td>
<td>Ser</td>
<td>Gly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Asn</td>
<td>Asn</td>
<td>Leu</td>
<td>Glu</td>
<td>His</td>
<td>Gly</td>
<td>Leu</td>
<td>Phe</td>
<td>Pro</td>
<td>Gln</td>
<td>Leu</td>
<td>Ser</td>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
</tr>
<tr>
<td>305</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
<td>Thr</td>
<td>Ala</td>
<td>His</td>
<td>Gly</td>
<td>Ser</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Asn</td>
<td>Val</td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Glu</td>
<td>Gln</td>
<td>Tyr</td>
<td>Gln</td>
<td>Gln</td>
<td>Leu</td>
<td>Arg</td>
<td>Glu</td>
<td>Ala</td>
<td>Ala</td>
<td>Thr</td>
<td>Glu</td>
<td>Ala</td>
<td>Glu</td>
<td>Lys</td>
</tr>
<tr>
<td>335</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Leu</td>
<td>Gln</td>
<td>Lys</td>
<td>Tyr</td>
<td>Ala</td>
<td>Glu</td>
<td>Ser</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Asp</td>
<td>His</td>
<td>Leu</td>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td>350</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>Gln</td>
<td>Glu</td>
<td>Lys</td>
<td>Lys</td>
<td>Ile</td>
<td>Leu</td>
<td>Lys</td>
<td>Asp</td>
<td>Phe</td>
<td>His</td>
<td>Gln</td>
<td>Lys</td>
<td>Lys</td>
<td>Asn</td>
</tr>
<tr>
<td>365</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Ile</td>
<td>Ser</td>
<td>Phe</td>
<td>Gln</td>
<td>Gln</td>
<td>Thr</td>
<td>Thr</td>
<td>Ala</td>
<td>Met</td>
<td>Val</td>
<td>Thr</td>
<td>Leu</td>
<td>Arg</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td>380</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Lys</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
<td>Ala</td>
<td>Ile</td>
<td>Thr</td>
<td>Ser</td>
<td>Thr</td>
<td>Ser</td>
<td>Ile</td>
<td>Leu</td>
</tr>
<tr>
<td>405</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Gly</td>
<td>Arg</td>
<td>Arg</td>
<td>Tyr</td>
<td>Asp</td>
<td>Asp</td>
<td>Asp</td>
<td>Asp</td>
<td>Ile</td>
<td>Phe</td>
<td>Pro</td>
<td>Phe</td>
<td>Pro</td>
<td>Gly</td>
<td>Pro</td>
</tr>
<tr>
<td>420</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Asn</td>
<td>Asp</td>
<td>Glu</td>
<td>Asn</td>
<td>Ser</td>
<td>Gly</td>
<td>Glu</td>
<td>Asn</td>
<td>Asp</td>
<td>Asp</td>
<td>Asp</td>
<td>Asp</td>
<td>Pro</td>
<td>Thr</td>
<td>Asp</td>
</tr>
<tr>
<td>435</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Gln</td>
<td>Asp</td>
<td>Thr</td>
<td>Thr</td>
<td>Pro</td>
<td>Thr</td>
<td>Ile</td>
<td>Asp</td>
<td>Val</td>
<td>Ile</td>
<td>Asp</td>
<td>Pro</td>
<td>Asn</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td>450</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Tyr</td>
<td>Asn</td>
<td>Tyr</td>
<td>Ser</td>
<td>Asp</td>
<td>Tyr</td>
<td>Ala</td>
<td>Asn</td>
<td>Asp</td>
<td>Ala</td>
<td>Lea</td>
<td>Ser</td>
<td>Ala</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>Leu</td>
<td>Val</td>
<td>Leu</td>
<td>Phe</td>
<td>Asp</td>
<td>Leu</td>
<td>Glu</td>
<td>Asp</td>
<td>Glu</td>
<td>Asp</td>
<td>Ala</td>
<td>Asp</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td>Gln</td>
<td>Asn</td>
<td>Thr</td>
<td>Pro</td>
<td>Glu</td>
<td>Lys</td>
<td>Asn</td>
<td>Asp</td>
<td>Arg</td>
<td>Pro</td>
<td>Ala</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Arg</td>
<td>Asn</td>
<td>Gly</td>
<td>Gln</td>
<td>Asp</td>
<td>Gln</td>
<td>Asp</td>
<td>Gly</td>
<td>Asn</td>
<td>Gln</td>
<td>Gln</td>
<td>Gly</td>
<td>Glu</td>
<td>Thr</td>
<td>Ala</td>
</tr>
</tbody>
</table>
Pro Arg Val Ala Pro Asn Gln Tyr Arg Asp Lys Pro Met Pro Gln Val
Gln Asp Arg Ser Gln Asn His Asp Gln Thr Leu Gln Thr Gln Ser Arg
Val Leu Thr Pro Ile Ser Gln Ala Asp Pro Ser Asp His Asn Asp
Gly Asp Asn Gln Ser Ile Pro Pro Leu Gln Ser Asp Asp Gln Gly Gly Ser
Thr Asp Thr Thr Ala Ala Gln Thr Lys Pro Ala Thr Ala Pro Pro Ala
Pro Val Tyr Arg Ser Ile Ser Val Asp Ser Val Pro Ser Glu Asn
Ile Pro Ala Gln Ser Asn Gln Thr Asn Asn Gln Asp Asn Val Arg Asn
Asn Ala Gln Ser Gln Ser Ile Ala Gln Met Tyr Gln His Ile Leu
Lys Thr Gln Gly Pro Phe Asp Ala Ile Leu Tyr Thr His Met Met Lys
Glu Glu Pro Ile Ile Phe Ser Thr Ser Asp Gly Lys Glu Tyr Thr Tyr
Pro Asp Ser Leu Glu Asp Tyr Pro Pro Trp Ser Leu Ser Gln Glu
Ala Met Asn Glu Asp Asn Arg Phe Ile Thr Met Asp Gly Gln Gln Phe
Tyr Trp Pro Val Met Asn His Arg Asn Lys Phe Met Ala Ile Leu Gln
His His Arg

<210> SEQ ID NO 4
<211> LENGTH: 373
<212> TYPE: PROTEIN
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus SGP viral protein
<400> SEQUENCE:
Met Val Thr Ser Gly Ile Leu Gln Leu Pro Arg Glu Arg Phe Arg Lys
  1   5   10   15
Thr Ser Phe Phe Val Trp Val Ile Leu Phe His Lys Val Phe Pro
  20  25  30
Ile Pro Leu Gly Val Val His Asn Thr Leu Gln Val Ser Asp Ile
  35  40  45
Asp Lys Leu Val Cys Arg Asp Lys Leu Ser Ser Thr Ser Gln Leu Lys
  50  55  60
Ser Val Gly Leu Asn Leu Glu Gly Asn Gly Val Ala Thr Asp Val Pro
  65  70  75  80
Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val
  85  90  95
Val Asn Tyr Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Asp
 100 105 110
Ile Lys Lys Ala Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly
Val Arg Gly Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
115 120 125
Gly Pro Cys Pro Glu Gly Tyr Ala Phe His Lys Glu Gly Ala Phe Phe
130 135 140
Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Ser Thr Thr Phe
145 150 155 160
Ser Glu Gly Val Val Ala Phe Leu Ile Leu Pro Glu Thr Lys Lys Asp
165 170 175 180 185 190
Phe Phe Gln Ser Pro Pro Leu His Glu Pro Ala Asn Met Thr Thr Thr Asp
190 195 200 205
Pro Ser Ser Tyr Tyr His Thr Val Thr Leu Asn Tyr Val Ala Asp Asn
210 215 220
Phe Gly Thr Asn Met Thr Asn Phe Leu Phe Gln Val Asp His Leu Thr
225 230 235 240
Tyr Val Gln Leu Glu Pro Arg Phe Thr Pro Gln Phe Leu Val Gln Leu
245 250 255
Asn Glu Thr Ile Tyr Thr Asn Gly Arg Arg Ser Asn Thr Thr Gly Thr
260 265 270
Leu Ile Trp Lys Val Asn Pro Thr Val Asp Thr Gly Val Gly Glu Trp
275 280 285
Ala Phe Trp Glu Asn Lys Lys Thr Ser Gln Lys Pro Phe Gln Val Lys
290 295 300
Ser Cys Leu Ser Tyr Leu Tyr Glu Pro Arg Ile Gln Ala Ala Thr
305 310 315 320
Arg Arg Arg Ser Leu Pro Pro Ala Ser Pro Thr Thr Lys Pro Pro
325 330 335
Arg Thr Thr Lys Thr Trp Phe Gln Arg Ile Pro Leu Gln Trp Phe Lys
340 345 350
Cys Glu Thr Ser Arg Gly Lys Thr Gln Cys Arg Pro His Pro Gln Thr
355 360 365
Gln Ser Pro Gln Leu
370
<210> SEQ ID NO 5
<211> LENGTH: 251
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus VP24 viral protein
<400> SEQUENCE: 5
Met Ala Lys Ala Thr Gly Arg Tyr Asn Leu Val Ser Pro Lys Lys Asp
1 5 10 15
Leu Glu Arg Gly Leu Val Leu Ser Asp Leu Cys Thr Phe Leu Val Asp
20 25 30
Gln Thr Ile Gln Gly Trp Arg Val Thr Trp Val Gly Ile Glu Phe Asp
35 40 45
Ile Ala Gln Lys Gly Met Ala Leu Leu His Arg Leu Lys Thr Ala Asp
50 55 60
Phe Ala Pro Ala Trp Ser Met Thr Arg Asn Leu Phe Pro His Leu Phe
65 70 75 80
<table>
<thead>
<tr>
<th>Gln</th>
<th>Asn</th>
<th>Ser</th>
<th>Asn</th>
<th>Ser</th>
<th>Thr</th>
<th>Ile</th>
<th>Glu</th>
<th>Ser</th>
<th>Pro</th>
<th>Leu</th>
<th>Trp</th>
<th>Ala</th>
<th>Leu</th>
<th>Arg</th>
<th>Val</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Gly</td>
<td>Ile</td>
<td>Glu</td>
<td>Asp</td>
<td>Gln</td>
<td>Leu</td>
<td>Ile</td>
<td>Asp</td>
<td>Gln</td>
<td>Ser</td>
<td>Leu</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Pro</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
<td>Ser</td>
<td>Leu</td>
<td>Val</td>
<td>Ser</td>
<td>Asp</td>
<td>Trp</td>
<td>Leu</td>
<td>Thr</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Thr</td>
<td>Asn</td>
<td>His</td>
<td>Phe</td>
<td>Gln</td>
<td>Met</td>
<td>Arg</td>
<td>Thr</td>
<td>Gln</td>
<td>His</td>
<td>Ala</td>
<td>Lys</td>
<td>Glu</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Lys</td>
<td>Met</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Val</td>
<td>Arg</td>
<td>Ser</td>
<td>Asn</td>
<td>Ile</td>
<td>Leu</td>
<td>Lys</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
<td>Asp</td>
<td>Ala</td>
<td>Leu</td>
<td>His</td>
<td>Val</td>
<td>Arg</td>
<td>Tyr</td>
<td>Asn</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td>Ile</td>
<td>Glu</td>
<td>Ile</td>
<td>Gly</td>
<td>Thr</td>
<td>Arg</td>
<td>Arg</td>
<td>Asn</td>
<td>His</td>
<td>Thr</td>
<td>Ile</td>
<td>Thr</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Met</td>
<td>Gly</td>
<td>Phe</td>
<td>Leu</td>
<td>Val</td>
<td>Glu</td>
<td>Leu</td>
<td>Gln</td>
<td>Glu</td>
<td>Gln</td>
<td>Pro</td>
<td>Asp</td>
<td>Lys</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Asn</td>
<td>Gln</td>
<td>Lys</td>
<td>Pro</td>
<td>Gly</td>
<td>Pro</td>
<td>Gly</td>
<td>Pro</td>
<td>Val</td>
<td>Lys</td>
<td>Phe</td>
<td>Leu</td>
<td>Leu</td>
<td>His</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Thr</td>
<td>Phe</td>
<td>Lys</td>
<td>Ala</td>
<td>Leu</td>
<td>Ile</td>
<td>Lys</td>
<td>Pro</td>
<td>Ala</td>
<td>Thr</td>
<td>Lys</td>
<td>Met</td>
<td>Gln</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ile</td>
<td>Leu</td>
<td>Glu</td>
<td>Phe</td>
<td>Asn</td>
<td>Ser</td>
<td>Ser</td>
<td>Leu</td>
<td>Ala</td>
<td>Ile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 6
<211> LENGTH: 289
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FRAGMENT:
<221> NAME/KEY: misc.feature
<222> OTHER INFORMATION: Bundibugyo ebolavirus VP30 viral protein
<400> SEQUENCE: 6

<table>
<thead>
<tr>
<th>Met</th>
<th>Asp</th>
<th>Ser</th>
<th>Phe</th>
<th>His</th>
<th>Glu</th>
<th>Arg</th>
<th>Arg</th>
<th>Ser</th>
<th>Arg</th>
<th>Arg</th>
<th>Thr</th>
<th>Ile</th>
<th>Arg</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Arg</td>
<td>Asp</td>
<td>Gly</td>
<td>Pro</td>
<td>Ser</td>
<td>His</td>
<td>Glu</td>
<td>Val</td>
<td>Arg</td>
<td>Thr</td>
<td>Arg</td>
<td>Ser</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Ser</td>
<td>His</td>
<td>Arg</td>
<td>Ser</td>
<td>Gly</td>
<td>Tyr</td>
<td>His</td>
<td>Thr</td>
<td>Pro</td>
<td>Arg</td>
<td>Ser</td>
<td>Ser</td>
<td>Ser</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Val</td>
<td>Pro</td>
<td>Thr</td>
<td>Val</td>
<td>Phe</td>
<td>His</td>
<td>Arg</td>
<td>Lys</td>
<td>Arg</td>
<td>Thr</td>
<td>Asp</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td>Ala</td>
<td>Pro</td>
<td>Lys</td>
<td>Asp</td>
<td>Ile</td>
<td>Cys</td>
<td>Pro</td>
<td>Thr</td>
<td>Leu</td>
<td>Arg</td>
<td>Arg</td>
<td>Gly</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Asp</td>
<td>Ser</td>
<td>Asn</td>
<td>Phe</td>
<td>Cys</td>
<td>Lys</td>
<td>Asp</td>
<td>His</td>
<td>Gln</td>
<td>Leu</td>
<td>Glu</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Ala</td>
<td>Arg</td>
<td>Lys</td>
<td>Thr</td>
<td>Cys</td>
<td>Gly</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Gln</td>
<td>Gln</td>
<td>Leu</td>
<td>Asn</td>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
<td>Pro</td>
<td>Lys</td>
<td>Arg</td>
<td>Thr</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ile</td>
<td>Ala</td>
<td>Asp</td>
<td>Asp</td>
<td>Phe</td>
<td>Gln</td>
<td>Gln</td>
<td>Lys</td>
<td>Asp</td>
<td>Gly</td>
<td>Pro</td>
<td>Lys</td>
<td>1le</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Thr</td>
<td>Leu</td>
<td>Glu</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
<td>Tyr</td>
<td>Thr</td>
<td>Ser</td>
<td>Lys</td>
<td>Gln</td>
<td>Asp</td>
<td>Ile</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ile</td>
<td>Asp</td>
<td>Asp</td>
<td>Ser</td>
<td>Arg</td>
<td>Leu</td>
<td>Arg</td>
<td>Ala</td>
<td>Leu</td>
<td>Leu</td>
<td>Thr</td>
<td>Leu</td>
<td>Cys</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Met Thr Arg Lys Phe Ser Lys Ser Gln Leu Ser Leu Leu Cys Glu Ser
180 185 190
His Leu Arg Arg Glu Gly Leu Gly Gln Asp Gln Ser Glu Ser Val Leu
195 200 205
Glu Val Tyr Gln Arg Leu His Ser Asp Lys Gly Gly Asn Phe Glu Ala
210 215 220
Ala Leu Trp Gln Gln Trp Asp Arg Gln Ser Leu Ile Met Phe Ile Thr
225 230 235 240
Ala Phe Leu Asn Ile Ala Leu Gln Leu Pro Cys Glu Ser Ser Ser Val
245 250 255
Val Ile Ser Gly Leu Arg Leu Val Pro Gin Ser Glu Asp Thr Glu
260 265 270
Thr Ser Thr Tyr Thr Glu Thr Arg Ala Trp Ser Glu Glu Gly Gly Pro
275 280 285
His

<210> SEQ ID NO 7
<211> LENGTH: 341
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<222> NAME/KEY: micr_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus VP35 viral protein
<400> SEQUENCE: 7
Met Thr Ser Asn Arg Ala Arg Val Thr Tyr Asn Pro Pro Pro Thr Thr
1 5 10 15
Thr Gly Thr Arg Ser Cys Gly Pro Glu Leu Ser Gly Trp Ile Ser Glu
20 25 30
Gln Leu Met Thr Gly Lys Ile Pro Ile Thr Asp Ile Phe Asn Glu Ile
35 40 45
Glu Thr Leu Pro Ser Ser Pro Ser Ser His Ser Lys Ile Lys Thr
50 55 60
Pro Ser Val Gln Thr Arg Ser Val Gin Thr Gin Thr Asp Pro Asn Cys
65 70 75 80
Asn His Asp Phe Ala Glu Val Val Lys Met Leu Thr Ser Leu Thr Leu
85 90 95
Val Val Gln Lys Gin Thr Leu Ala Thr Glu Ser Leu Glu Gin Arg Ile
100 105 110
Thr Asp Leu Glu Gly Ser Leu Lys Pro Val Ser Glu Ile Thr Lys Ile
115 120 125
Val Ser Ala Leu Asn Arg Ser Cys Ala Glu Met Val Ala Lys Tyr Asp
130 135 140
Leu Leu Val Met Thr Thr Gly Arg Ala Thr Ala Thr Ala Ala Ala Thr
145 150 155 160
Glu Ala Tyr Trp Ala Glu His Gly Arg Pro Pro Pro Gly Pro Ser Leu
165 170 175
Tyr Glu Glu Asp Ala Ile Arg Thr Lys Ile Gly Lys Gin Gly Asp Met
180 185 190
Val Pro Lys Glu Val Gin Glu Ala Phe Arg Asn Leu Asp Ser Thr Ala
195 200 205
Leu Leu Thr Glu Glu Asn Phe Gly Lys Pro Asp Ile Ser Ala Lys Asp
210 215 220
Leu Arg Asn Ile Met Tyr Asp His Leu Pro Gly Phe Gly Thr Ala Phe
225 230 235 240
His Gln Leu Val Gln Val Ile Cys Lys Leu Gly Lys Asp Asn Ser Ser
245 250 255
Leu Asp Val Ile His Ala Glu Phe Gln Ala Ser Leu Ala Glu Gly Asp
260 265 270
Ser Pro Gln Cys Ala Leu Ile Gln Ile Thr Arg Ile Pro Ile Phe
275 280 285
Gln Asp Ala Ala Pro Pro Val Ile His Ile Arg Ser Arg Gly Asp Ile
290 295 300
Pro Lys Ala Cys Gln Lys Ser Leu Arg Pro Val Pro Pro Ser Pro Lys
305 310 315 320
Ile Asp Arg Gly Trp Val Cys Ile Phe Gin Leu Gin Asp Gly Lys Thr
325 330 335
Leu Gly Leu Lys Ile
340

<210> SEQ ID NO 8
<211> LENGTH: 326
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus VP40 viral protein
<400> SEQUENCE: 8
Met Arg Arg Ala Ile Leu Pro Thr Ala Pro Pro Glu Tyr Ile Glu Ala
1 5 10 15
Val Tyr Pro Met Arg Thr Val Ser Thr Ser Ile Asn Ser Thr Ala Ser
20 25 30
Gly Pro Asn Phe Pro Ala Pro Asp Val Met Ser Asp Thr Pro Ser
35 40 45
Asn Ser Leu Arg Pro Ile Ala Asp Asn Ile Asp His Pro Ser His
50 55 60
Thr Pro Thr Ser Val Ser Ala Phe Ile Leu Glu Ala Met Val Asn
65 70 75 80
Val Ile Ser Gly Pro Lys Val Leu Met Lys Gin Ile Pro Ile Trp Leu
85 90 95
Pro Leu Gly Val Ala Asp Gin Lys Thr Tyr Ser Phe Asp Ser Thr Thr
100 105 110
Ala Ala Ile Met Leu Ala Ser Tyr Thr Ile Thr His Phe Gly Lys Thr
115 120 125
Ser Asn Pro Leu Val Arg Ile Asn Arg Leu Gly Pro Gly Ile Pro Asp
130 135 140
His Pro Leu Arg Leu Arg Ile Gly Asn Gin Ala Phe Leu Gin Glu
145 150 155 160
Phe Val Leu Pro Pro Val Gin Leu Pro Gin Tyr Phe Thr Phe Asp Leu
165 170 175
Thr Ala Leu Lys Leu Ile Thr Gin Pro Leu Pro Ala Ala Thr Thr Thr
180 185 190
Asp Asp Thr Pro Thr Gly Pro Thr Gly Ile Leu Arg Pro Gly Ile Ser
195 200 205
Phe His Pro Lys Leu Arg Pro Ile Leu Leu Pro Gly Lys Thr Gly Lys

-continued
<210> SEQ ID NO: 9
<211> LENGTH: 676
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE: misc_feature
<223> OTHER INFORMATION: Bundibugyo ebolavirus GP viral protein

<400> SEQUENCE: 9

Met Val Thr Ser Gly Ile Leu GlN Leu Pro Arg Glu Arg Phe Arg Lys
1 5 10 15
Thr Ser Phe Phe Val Trp Val Ile Leu Phe His Lys Val Phe Pro
20 25
Ile Pro Leu Gly Val Val His Asn Thr Leu GlN Val Ser Asp Ile
35 40 45
Asp Lys Leu Val Cys Arg Asp Lys Leu Ser Ser Thr Ser GlN Leu Lys
50 55 60
Ser Val Gly Leu Asn Leu Glu Gly Asn Gly Val Ala Thr Asp Val Pro
65 70 75 80
Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val
85 90 95
Val Asn Tyr Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Asp
100 105 110
Ile Lys Lys Ala Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly
115 120 125
Val Arg Gly Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
130 135 140
Gly Pro Cys Pro Glu Gly Tyr Ala Phe His Lys Glu Gly Ala Phe Phe
145 150 155 160
Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Ser Thr Thr Phe
165 170 175
Ser Glu Gly Val Val Ala Phe Leu Ile Leu Pro Glu Thr Lys Asp
180 185 190
Phe Phe GlN Ser Pro Pro Leu His Glu Pro Ala Asn Met Thr Thr Asp
195 200 205
Pro Ser Ser Tyr Tyr His Thr Val Thr Leu Asn Tyr Val Ala Asp Ann
210 215 220
-continued

Phe Gly Thr Asn Met Thr Asn Phe Leu Phe Glu Val Asp His Leu Thr
225 230 235 240

Tyr Val Gln Leu Glu Pro Arg Phe Thr Pro Glu Phe Leu Val Glu Leu
245 250 255

Asn Glu Thr Ile Tyr Thr Asn Gly Arg Arg Ser Asn Thr Thr Gly Thr
260 265 270

Leu Ile Trp Lys Val Asn Pro Thr Val Asp Thr Gly Val Gly Glu Trp
275 280 285

Ala Phe Trp Glu Asn Lys Lys Asn Phe Thr Lys Thr Leu Ser Ser Glu
290 295 300

Glu Leu Ser Val Ile Phe Val Pro Arg Ala Glu Asp Pro Gly Ser Asn
305 310 315 320

Gln Lys Thr Val Thr Pro Thr Ser Phe Ala Asn Asn Glu Thr Ser
325 330 335

Lys Asn His Glu Asp Leu Val Pro Glu Asp Pro Ala Ser Val Val Glu
340 345 350

Val Arg Asp Leu Glu Arg Glu Asn Thr Val Pro Thr Pro Pro Pro Asp
355 360 365

Thr Val Pro Thr Thr Leu Ile Pro Asp Thr Met Glu Glu Glu Thr Thr
370 375 380

Ser His Tyr Glu Pro Asn Ile Ser Arg Asn His Glu Arg Asn
385 390 395 400

Asn Thr Ala His Pro Glu Thr Leu Ala Asn Asn Pro Asp Asn Thr
405 410 415

Thr Pro Ser Thr Pro Pro Glu Asp Gly Glu Arg Thr Ser Ser His Thr
420 425 430

Thr Pro Ser Pro Arg Pro Val Pro Thr Ser Ile His Pro Thr Thr
425 430 440 445

Arg Glu Thr His Ile Pro Thr Thr Met Thr Ser Ser His Asp Thr Asp
450 455 460

Ser Asn Arg Pro Asn Ile Ser Glu Ser Thr Glu Pro Gly
465 470 475 480

Pro Leu Thr Asn Thr Thr Arg Gly Ala Ala Asn Leu Leu Thr Gly Ser
485 490 495

Arg Arg Thr Arg Arg Glu Ile Thr Leu Arg Thr Gln Ala Lys Cys Asn
500 505 510

Pro Asn Leu His Tyr Trp Thr Thr Gln Asp Glu Gly Ala Ala Ile Gly
515 520 525

Leu Ala Trp Ile Pro Tyr Phe Gly Pro Ala Ala Glu Gly Ile Tyr Thr
530 535 540

Glu Gly Ile Met His Asn Glu Asn Gly Leu Ile Cys Gly Leu Arg Gln
545 550 555 560

Leu Ala Asn Glu Thr Thr Gln Ala Leu Gin Leu Phe Leu Arg Ala Thr
565 570 575

Thr Glu Leu Asp Thr Phe Ser Ile Leu Asn Arg Lys Ala Ile Asp Phe
580 585 590

Leu Leu Gin Arg Trp Gly Gly Thr Cys His Ile Leu Gly Pro Asp Cys
595 600 605

Cys Ile Glu Pro His Asp Trp Thr Lys Asn Thr Asp Lys Ile Asp
610 615 620

Gln Ile Ile His Asp Phe Ile Asp Lys Pro Leu Pro Asp Gln Thr Asp
<210> SEQ ID NO 10
<211> LENGTH: 18935
<212> TYPE: DNA
<213> ORGANISM: Cote divoire ebolavirus

<400> SEQUENCE: 10

cggacacaca aaaaagaaga aggtttttttg atcttttatg tgctggagaata acatagagga  60
agattaataa ttttctctct attgacacctt acattaagat taagattcct attgatctgt  120
tacttactct gaggataata attgggtgtc aagaatcccc cattcoccacag tggggggcaaa  180
gagccgccca aagacatcc tttgctctatt caactaactt gttttgtctc aagtggccac  240
atatggataca taccagagag ttgacctaata tttcacaagct tagaggtgga cctagtgatat  300
cctcgggggt gttcataaat gagcoccattta acacacacacag agocctactg tttcctctgt  360
ttgccgtgca tttagatccta gagacaacttt aacccctcaca cttggccaaag cagggaataa  420
tctaacacaga ccggcgacgtg ttcttcagcct ttctagcccc aagatgggaga ctcgggcccc  480
caaagcagtg atgacacagca cggcagcagg ttctgaacacag attaccatata agacttttaac  540
agcgagatgg tccgcccaac aagggctcttg gagcacaaggtgt gtcacacagc ttcacccag  600
tacaaaccta gaagaaatag gcaatctgct cacctcaacgct tttgagctgt tgtgtgatt  660
tcaagaggtgt gcagacacgt ttccctgtat gctatggtaa ctcctcatcgt ctacagggta  720
cuaacacaa tttctttgacc cagatcgact caagaccttt gagggtctagtg gttctcgttt  780
tggagtggg ccaaaaaagag gagaagcg ctcgagcaccctaa cattcgccc tggctcctagt cggcttccag  840
tggcaagac atccgagggag acctgctgac aagcgtcgaa gaggagacac cagacagaa  900
tgcgcagcag ttcctctct ttctttctgt attccctttg aagctagttg tggagaaaaa  960
agcccgtgtca aaaaagggctg aggcgggata tcaggtcataa tctgccaggg gattgcatca 1020
atgccccaca gccgtggcagt cagttggaca catgatgttcat cacaagccct gtttatgactga 1080
aataatataa ttaagtccct cccctattac taacaagagct catatgtgat caggacagca 1140
tgtaagacgt ctcgacagc aaactccttg agctacacag gttttttcgt gatttatgatt 1200
cgttaaaaac gctgctagac acatcctttca gaaacacag cacagagtgc gttctcttcct 1260
ttggtcgcaga actcctcataa cagaaaaac gtttttgtcg ctctctacgct 1320
gctagccaa cttgggaggt atgccctttc tgtcctggttg tttgctcttt ctgagctcag 1380
caactctgag caagctcgtcg tttcctgcctc tctgaccatt gcttcttggt tgcgaagcc 1440
acacggcggtt acctcggccat ggtttaagtt tggggacacac tgtcagcacc tacagagaa 1500
agcactctgac gacaggacacaatcagctgaaa aactgcagatcgcttcgcag tggacctct 1560
agttctcgt ctacacagga aagatgcatt gaaacacctt cttcagaaaga aaaaatgat 1620
cagctttcgc cagacacacg cattgctcag acctgcaaga gaaagctctg ccaagctcag 1680
tggagccata cttcttcctc cgcagcagaa cagtttatggt aagcacaagca 1740
-continued

tacccccctt ctcttgccca tcaatgataa cgaagaactca gaacacgcaag aogatgacc 1800
aaacgattct cagagacact ccatccctga tcatatgatt gaccgcggag atgpcaga 1860
cacaatatt ggagaacatct ctatgtgagac ggcaagagc cctgaagacc ttgtcttttt 1920
tgaccctgca gatctgagcc agagatgac ccagccgctca aagtgcacag agaaccacc 1980
cacaacagct tctccagga acgacagataa caaacaaggt acatggaacc gaaacccgac 2040
taatatgcca aagaagact ccacacaaad ccaatgcaaat cctgcaacag gggctcaaga 2100
atcgcacag ggtaacactc aggataaccc aacaccccat cagatctcaaa cttccacactcag 2160
cgasaagacc ggctccactg ttctcaaatgag atgacacctg gattgacatc cttcctttgga 2220
atcagacaag aaaaaaccagg atgacacaaac ccattacacc aaaaaaaaata ccactgactc 2280
acacagacctt gtttacgga gtaaatgacaa aagagagccg gctccgcaag aaaaactccca 2340
gagaaccacc aacacgacggt gttgctagta ctaaaccgaa aacaaactcct cttcgcagca 2400
atcagttgca gaaatgcatac gacacatcct ccaacacaag ggacacctgctt atgacattc 2460
atgatatatc atgacagaggt ggcggccgat ttgctttgtcac aagagatgctg 2520
cgtataccct gatccctctg aagggagcaca ttccacggtc ctcagcggaa aaaagccctt 2580
gattgcagcc aaagtggttta ccaaatgcttg tgcgaacaca tcctactggcg cgtgaatgac 2640
tcagagcactc aatctctcagtc ctacccctca gcaacacacaa taattttctcataa tattgacag 2700
atcattgtca ggttatcactt ccatttccttcg caaaccagcc tgaaaaaaccag ctcacacacc 2760
gccctacac aggatacctt ggagaacact caacagactc agacgagtct gcaaccacc 2820
cgagccgctc caccacccacc acacacacacc ataatcttcca cgaacaccacc gcacaacactc 2880
agataccac aacactcctc acaacacccg actccgacag tccacacacc aaccacacac ctcacacacc 2940
agacacgcg cccacccacat ccagagaccc cgctggcccg gactacactgg cagcttccag 3000
aatggttta taagaaaaa tatatatatacag aattaaaaaat cttcatcaca attgacacaga 3060
cttattcct ctggtattc ttcatttgta ctcaacaaaaag ggcctcaggg ggaatgatttt 3120
ccactgacc tggcggcatc aatgtccctt cttaccaact cagagacaag ctgtaacagttg 3180
gccctgact caactgcgtgg atctcgcgaac ataaaaagct ccgctgctcag 3240
aataatcctaga ggcactaggg cccacacataa gtoaggtggt ccagagcgctt cccagacacta 3300
aaaaacgcgc cccacggcact gcgaacaccc agacacagcc gctatcgggt tgaataccaca 3360
attttgaga cggtaacaca gcctacactc tttacacaccat cggctacacc caaacaagctc 3420
cttaactgta gtcctcggaa ccaagcatca tagatcttga gaatgcctta aagcccaattg 3480
atgacatggc taagatcttct tctgacatttga atagatctttg tggctgagatg cagagaaaaat 3540
atgatcttct ctggtgacacatac ggagccggcg caacgcggac ccggctgcgtc aatcagggctc 3600
attggggaga ccagactgacc ccacacactc gaccaatct tttagaagag aggtgctta 3660
gagacagaa agtaacacca aagctttggc aagcgcggat atatcgtgcac 3720
agagacactc aacagctactc gatgagccga ctaacacggag aacatcttgg aagcgcggat 3780
atgtgggagt ctaacacggag aacatcttgg aagcgcggat atatcgtgcac 3840
tggtgcaggt atatttggaag ctatccattgc acatgacatt atatcgtgcac 3900
agtgcacccg cccagctctc gagatgtgtt ttgccctgcac atatcgtgcac 3960
aaggggccag ccccccacgt ccgcttggtg cttcgcaggc gtcctgcac 4020
-continued

tttccacaaa ggtggttaat tgcgaaagtg gagaattggc tgggaacctg tataacctgg 6360
cataaaaga aggtaatcgt agtgagttgg taccagaga cccagggaga gttgaggaatt 6420
tttcccgttg cgccgtatga cacaaagct caggaacttg aacattgcca gagaagacatg 6480
cctttccaa aaggaggtac ttctctctgt atgaccccat ccgcatcaaca atctttttac 6540
gggattacaac ctttccggaa ggaagttattg cttctctctg attcctaaag ggggaaaggg 6600
atattttccta gttttcttca tggtattgag ctggtacact gacacagagat cccgccatgt 6660
acttacacca gacaacaata aaccagctgg ttgataaatt tgggaaccaac accacagagt 6720
tctctgtcctt aacagcttat tggagctagt tgcgacgctg aaccagactt aacaccaaat 6780
tttctgtctt cotaatgaac aacatcctaat ctgataaacc cagaagtaac aacacagga 6840
aaccatctct gaaaaaatat cccacctgttg atacagccag gttgtgagttgg gttttcttgg 6900
aataaaaaaa aaccaccaaa aacocctcct aacatggaag tggctttctgg tacctgtgacc 6960
agaaaccccg aaccagctcc tgttaacagac aggcagcttg tttctctcao acctccgcca 7020
ccccccgcca ggcaggaacct acaaaatggt gtttccgagatt ccctcctccct cagttgctca 7080
gctgcaaaat accaagggaa aagcataaat gcaccaacta gtcagcgggtg tatcaaccaac 7140
caacccctct ccctttctcct ctaattgtcog caaactctgt catacacaaat catttatcgg 7200
cctttcccgg cccacagcgac accaagccac accacgctcc gccaaagcacc cccacacaaac 7260
aaccacagcc aaccagcttc ccgcagctaa cccacactca ggcocctcca ctagacccgc 7320
ggggacatcc gcccacaggag tttccacaca cacagagaac caaggccaga ttgagtaagac 7380
aaccacacac accttcccag aacagcacac tgcgctccgt gctccacacac ggcctctga 7440
ccccgcagaa ctgctttgca accttttcttg gcagacacaa atacagggggt tttaacaaatct 7500
ccttgaaggg ttgagttcag caacacacat gtaagctggc tattcagcttt ggttattgatt 7560
ccccacacta cacatgtgag cagctcctgg tggaggtgtg gcaatagttt tagcttggat 7620
accacaacttc ggcacagcttg ctgagggatatt ctaacagagatt gaatccaaaaa 7680
ttgggattgct tggggattgct gggagtggag ccaaaaagag cacaagactt tttaattttgct 7740
ctttgagcct agttcttgag gcagagcgtga cctgaaagac aacacactat ctgatattttct 7800
ccttgctcag acatgttgag gaaatgtgca cttttttggtt ttgatgtggtt gaattgaacc 7860
ccccacaggcg accaaactta tcaggtgaaa aatacttgca aataatgtct tattttagctt 7920
taatataatct ccacacctag attggacccc caacaatggt aacatgattt 7980
tttctcctg gtaagtagttt attggatattttt atggaatagg accctcacat attcccgggg 8040
ciaacatctg ccacacacta atatactgct atatacctcta atatcccctc aataataatc 8100
tggctttggag aggttaatatg tttatatattttt attgagatag acccttacca 8160
taatattttt tattgctagtt ccaagaatgtcatcagttt catctactac caaatatatc 8220
atggaggaacc atgcagctttg acctacacat tggctttcgtt tcttaactac ctaacagcac 8280
acagtaagca aagctccata ttatgattta ctgtatcttt atcttataaaat acatgattttt 8340	taacaaacc gacatggtcag gcctctagg atggtgacatc gcaaccacaa ataatccttc 8400
ttggagactc ttcagcctta aaaaatatgc tttgaggttg gttctcttgat 8460
aacacataa attgttttta cccacacacta aacagcggctg acctatgtcctg tttgcttgg 8520
agacgtgcct ccgaggtcct ggcacaaacg acaagaaaagttg cagctcctca tttatgtcgg 8580
gagagagccc cctctcagc ttagtatgca aagtaatacg ataacacaagg gagtgcctcg 9640
cagatcctg gcccacatgt ctctcctcgg aaaaagcag atatttggc agttccacca 9700
gccctaaag atgtattggc gacttttaaa aagagcttcc tttgtgcaag caaatttcgt 9760
aaaaagata cccccactgaa aagtaatacg gtagagaggt taacattgtg gattgcctcg 9820
aagacatgtg gctccacagga acaacacacta agcatagtgg ctctaaagaaggctcctg 9880
gctacttcc tttgtgagga tttccaaacaa aagatggggc ttaagtttac atctgctgatg 9940
cattatagga cagacagata tttgctccaa cggacacta aagacacagta gatttcaaga 9000
ttaagactt tattacccct tttgctctgt attaaggcga aatcctctaa atttcctctt 9060
agttgtcat gtgaaggcct ctacaagccga ggacatttcg gctagaacac acaacaagta 9120
gttgtggagg tataactccg ctccacacag gataagggg tgaatttggc ggcgacatca 9180
tggcagcgct ggtgctggcga acaacatgcc atctcatgta atgcctcataa cagctttttt 9240
atatatatcc tattcgagag tttcactgtt gttatttccg gttttagaat gttatcatacc 9300
cattctgccg ccttctcagtg ttaacctccgg tccgaaccct gcaacctgctc gaaagagggga 9360
agttcccaat gaaagcccaaa atccaaaggg cgcctaaaaa atcccttttg acaacatcata 9420
acatacata caatacttggaa gcgattggaa taatgtgtga ttctgaaggag attagttgtt 9480
gcggctcacc taacagcagc gcaatataaa tctactctac agcagtttaa tttctaatat 9540
aagatggtaa aacagggctc gggttccagta ttgcttccccca ggtctcactt ctcttgtttta 9600
aatgttagaa gtagcaacaa gttgacgttaa taattgataac tattatatccg aacagcaaca 9660
acctatatc aagatgttgtt atatcctggca atctctaaaa taatgtctac atgttgctta 9720
caatatgaa aagacacagt gaccaatcaca attagtcca taaocgact aacctcata 9780
tagaacaagc agaaccgcgc acagacgcag caataacaca tgaaccacaa tattgtaactg 9840
tataatttt caattcgtaaa tccaaacaaa atattagaaaa gtaactgca gtttgaaaa 9900
ataacatcat catgcctcag gtttataat gggcttcgta gaaatactctt atccctctca 9960
acotaotca aagaagttaa ttcogctacag tggggactt gaaogaaacaa ggaagcaacc 10020
cttttccacg taacagcagctgtgctggggc acagacggta ttaacctgcag acgcgcacat 10080
tccctataat tcaagcggta acagatcaaa aaccaagcagc tattagacac tgtctgattaa 10140
tagttacgca taagctcaat attttataa acaatggcggaa tctatatcttc caaacaagaat 10200
atttctcagtt taacagggagc aacacagctc gcacaggagc acaactgccgt acaaatcata 10260
agaaaaagac aacccctgac ccagagttca aagatacata cttaacagga taaatcctgg 10320
gctctgctgg agttccacaa tggcaaggtc atctgaggtg acaaccttat cttccaaaag 10380
aaaagatotg aaaaaagcttg gttctctaat gacottttcga atctctcgag ggcagcagac 10440
gttgggtgg ctaattgctac tgggtggtggc attgaatttg aagggagaccac gaaaggcttg 10500
gcctatttgc acagcactcc gacoggctag tttgcttcag cctggctcaat gacogcgaaac 10560
taatctccaa atctctttttt ccacctccgaaac atactctttgtg aacggcacaat 10620
tgggtgcagt cacgtagctg ttggtgctcota taatactgct gatogaaacc 10680
ctgcttcgact cagcagctacg tatctactggt gcacttgcctactgtagggg 10740
caatcgcgca cagcagcaggt taaagggcata atccgctcataa aatagttggc cctggtgcga 10800
tcacaatacct aagacatcagt attgtgcggat gtggctgaa ttcaaatgga 10860
ctttctcagta gcaatgaaat tggaacccaa aagccatatca ttataattac cggga caaat 10920
atgggttttt tcggtaggtt gcaagacgct gcaaatcag ccaagcacs caggaaacca 10980
ggccacgcca aattocctct ctctactgaa tcaactcgtg aagacaacctg taaaatcact 11040
ggcacccgaga tcggcagcat tattttggaa tccatagtt ttccttcgat ttaaccacac 11100
tcataacaat gtaaatctgt gatocattaa tcgcaocota gaocctttgat agaatctacta 11160
actattgacg atgtttcttg acatagggag aagaaactctg cccattagat agaggggttc 11220
cacattaat cccacacata cccatcatgt taacctcagt tagataccctc aagtatatc 11280
agctcactct accccctgcga ttgctactct aatcataata cctgataaca tttaatt 11340
tagtagatgt cttctatatt tagaagtttt tgtagaagta attgacatct caacttgatta 11400
cacataaca agaggtgtaa ttcgatagtt gttgaaata tttgtgcaaa taactgaact 11460
aaagagatga gatgggttaga agcttggagga agataaagaa aatgtgcgtc gggtgctttt 11520
cogctgtagaa gggaacacag acataatct tctctttttat acaacatggc tcaacacat 11580
acgccaatac tccagcnaag gttattccata cccatagtttt tagatacagt tgccttctgc 11640
actctgcct ct ggagtgtaa ttcgctacat ttttaaaata cccactaca aacactgtaa 11700
cacaaggaat atataacatt ctaaanaatat gagacaacctg ttcaacctgt ttgatgat 11760
gttgctttag caacacgacg aagggattttt ttgatctact cttctctctag tggctactca 11820
ggagagtgatct ctgtccctaa tgttatcataa aaatgacacta aattgatcgaat 11880
tatacttact aagatattccg otocrocacaa tatatactca atgagacggg agtgcaatac 11940
gatcactgtgg ataggagttt ggcaacaaaa atggccactc tatttccgta cggagtgtgtt 12000
ttccataaa tgtttcactg gttgatatct gcaattctag caotgtaggg aagcgatataa 12060
agagggaaata atctccacacag atgtgttcgc aagtgaaaat tggatgcagct cctagtttat 12120
ggagattata tttttttggaa aatacataa ttaacatac aacggtgataa aacagaccco 12180
ccacatcgag ccaagcaccggt gtaatcagaa cctgtgttca aagaggtgat tcaagcctcat 12240
acacacacatg gtcctatcct tacaagagat gtattoataa gttgttgaag aatacttaacc 12300
tgctgatta atataactcact gattgtgtct gttgccatct tagaggcttc atccctccta 12360
gattacctct taagacaaag agtgctgac otataacaag aagagattta ttaatactca 12420	ttgctataggt cagaggtta ccaagcataa aataactctgt agcgagttcg ttcgcgaaat 12480
atccaaactc gtccacatttc ctcctgagg aanaagagat ttcctactca atatacttta 12540
tgctgataa atataactccag gaaacttcga gggcoaccag aataaggagc acaacagatt 12600
cggagggtaa gggaacttcg tcaaatgctgt ataaaactta aagcaactac tcaacaactc 12660
tgtgagtgttc tttctcgtcga aagactctgcgg ggcacocctgt ttcctctgcag cggaaagctt 12720
ataccaaaaa taaacacggtg tctggtcact taaatcactg cttgtgccc aataatatctt 12780
aacatacattgt gttgtttctaa atacagactt gcaaaacttt atttgtgatag tcaaggtdag 12840
tggtagagtct ttggctggta gacatctta aacacgggca cttctcctta catcggaga 12900
aacacactt tccacactact tatggataaa aacocctttt ggaatcttta acacactcct 12960
ctctctctgt ttccttccac caaatgtggt aatgattttgctacatcattta taagactcgt 13020
gctctgtcgg tggcggacaag atcgtggcgc gcgttttttt ggacaaactg ttcctttgtt 13080
aacacacgca aataactcttcg tacaadaggg gtacctgcg aatctccttg aacagggact 13140
-continued

agtattcttg ctaactaggct agttgttccc acacaccactct ctcggaatt ttcaggagga 15480
ggtcagtccg caagagactg taattaaaaa tccgaagatg taattaattt ctctggttgc 15540
cctttgtgac taaagttagg gaaogtggct acttcttcctaa ttaacacca aacgggctc 15600
cctccattgt gaaagtctgg cagcgagagag gctcagcgcctt ataataggtt ttatactca 15660
acattgcat tggacactac aacoagtctgc gataattagt tgaattagtta gacaatcaca 15720
ttaagagctgt gttaaaaattct ctaaccttcct cttgataaattc cgcttttctca gagccagaag 15780
cctaaacat gggagagaata cttgtatgagt ctaaccctatt agagacctg tggagacgt 15840
aaacacttgta ttcaactcct aattctgcag acacacacga ccatacgagtt cccattaagcgt 15900
agtgggggaa caagatgacac ccacactccac ttcctggcat atataaagat agagactacta 15960
atatatttttg gtgcacactc cagttattatt ctaggccacaaccttatag aacacaaaaaa 16020
ttgcctttca atacaccaaat atatactcta gtaacactaaat ctaanatattt aacttccctg 16080
togttgaaga ccctttaaacc tactttgaaga cctctgtgctg ttctctcctg attaaataagt 16140
atgacactct acctctcaaat tttattagga ggacacgtctgtgtcaagtg accttcgcat 16200
gagccagaagt gtgtttcatct aactgcactct actgtctccct ctcactacgtc tagaaagtggt 16260
atattgcgaag gcacagacctgc tatctccactg gtaacactcaaat gacacaaaaaa 16320
ccattgcgtg tcacaggttt ttctacacac gctactgcat ttgtgcaata gttgcctttcc 16380
cagctgatag ttgtgtgtgc agaagcccccag actgctagttg tagacattgg aaatattgg 16440
atatgtgcta aagacagacgt agaagctcttt tttctgtcct cttctgtca gaaaaaaaaa 16500
cgatctaaaa atcagaccaaa aagagatgct aaaaaccttat ctcgctttgct gcacagccga 16560
aagaaaaatt caatcgctct tacaagatc oocagaaagcc cctgctctct ttgctgcctc 16620
cagacagcttc tctccacaccc tccccattgt cagcgggctgactacacgt tagaagagaca 16680
gttttgaagg cctccagacag tggcgattct ttccagcata caaccaacaac caacacacaac 16740
ccaacagtt gcctgcaagt atctcccaaat atcaagagcgc cagatggagtt ccctcccctgc 16800
aatgaggttgt tcaacagcaaa taacaaatatt ttgctctgct agaagatgtag tgggatcttt 16860
cttcattgtc ctctttaaat tgaagagaga cctctcctat cagcgtgggg aacagtgaca 16920
 gagttgtaaa ggtcattct gcagcgtggg acataacag ccccaaacaat atattgcgcg 16980
ttcaccggaaaa taattctccct aagcgactata ctctctctttt agtcattctta ggtattttaa 17040
aatatttaag tctctataact acacctgcsag gttgaggttg cggcttgactt actaccttatta 17100
caaaaaatt aatagtaaaag cttgttctcaatt aatartatct caacagagact ctctccatat 17160
geagaaatta tttctggaatt aactaacca aacagctttt ctccttaattt gctatttgctc 17220
ctgtggtgaa aataaaaaagt ctatataaacc agattacaag ggcacagttc cgtatattga 17280
aatccaaagtt ggttggccag ccacaaacct aggacccctgtagaagtaggctgtttcatc 17340
atggagtctg cacaacacaata ggctcagatattcagtag cgggttaactgtgc 17400
cctgtggtct cccatattga tcggaaagtc ctaaagacttg cgggttctcag aacctttctta 17460
agtcctagct atggaactct actggtcaag gataacttta cccctttggtt ccagcttggt 17520
tacttgatca agcggacactc cctcggccaa aacagttaagtg gtgggtttcct cttgctctca 17580
aacttctctct tcaatcccaag caaatcaacgt actcagatct aactatctctgtgctaatct 17640
atccaaacag cactctgact cacaacctag aggacgctcat tggggttagg caacccctgctc 17700
cagtatgcca atcataatgg gcatttagat tatattaatcc tccggttccc tccattggag 17760
aggyttttat accatagata caatattgct gattttcaga aagggcccttt gacctttatt 17820
gtccaaactc tagcgcacact gcagaccagg atagggaggt tggttaatgga ctataatc 17880
csasgacaas gcgcacacca acatatacat ttccaataaa ctaaamagg tgcatttaca 17940
aatgttgaas atgatgattt taaatctttt ctaaaatata aagccttaaa gcacattgc 18000
acagtgccag aggaactaag aggctcctcca gatccataa gtgtgtgcac tcgattctat 18060
catactgaa acgtgttcatt tcgaacaggg tttctctagac agactttata cttattcagc 18120
atgcagagtt cgagcaactaa actataatag aagttggtcag gcctttcttag tctatgc 18180
atagggtgatt ttcggttaga acctttggac tcaacactcc acaatgtagt acaatggtac 18240
caggacacta taagtaaatc gacccaataa aagagaattag gcacacacag agtttttcaag 18300
tgacaccccc tccctctcctt tattggtggt tgaattttctt atataccgatt gttaccccg 18360
gatataact tcacgttaat ataagaaatc acttttttgc actcacaatttt atcctaaat 18420
cacatcata cacgtccttacc ctcatatatt ataaagaaaa agttttctctt tttgctttta 18480
ttatataactt aatattttcg cagttttcaat tttactaaaaa tttaaattctg gttttttta 18540
aaggttatatt ttcgacacta acatatattt ttcattgttaa accatataag caaogcatga 18600
tgcgccttcc tctattcttt tctgttggtca cgcatttctc gcattttctt gatccccctca 18660
aacacttat ctggtctttc cgcgtctttc ttcgatctca gtttattttt ttaaaatctt 18720
aatcgggggtt atgtctgtat aaactaataa taccgaatta gttgtgattt aaaataaacg 18780	ttaaagattt tttttttctc cgcgtttttt agtatattag gaaagtctgca tcaagtttgc 18840
cctactccg aaaaagaaaa agggagcccc tattggtcatt ttaacgtgac aacaataat 18900
cctttacaaaa ttgcttttacttttttgtg gtcaca 18935

</210> SEQ ID NO 11
</211> LENGTH: 739
</212> TYPE: PRT
</213> ORGANISM: Bundibugyo ebolavirus
</220> FEATURE: misc_feature
</223> OTHER INFORMATION: Cote d'Ivoire ebolavirus NP protein
</400> SEQUENCE: 11

Met Glu Ser Arg Ala His Lys Ala Ala His Thr Ala Ser Gly

<table>
<thead>
<tr>
<th></th>
<th>Met</th>
<th>Glu</th>
<th>Ser</th>
<th>Arg</th>
<th>Ala</th>
<th>His</th>
<th>Lys</th>
<th>Ala</th>
<th>Ala</th>
<th>Ser</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Phe</td>
<td>Thr</td>
<td>Arg</td>
<td>Tyr</td>
<td>His</td>
<td>Lys</td>
<td>Ile</td>
<td>Leu</td>
<td>Thr</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Gln</td>
<td>Gly</td>
<td>Ile</td>
<td>Val</td>
<td>Arg</td>
<td>Gln</td>
<td>Arg</td>
<td>Val</td>
<td>Ile</td>
<td>Gln</td>
<td>Val</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Leu</td>
<td>Glu</td>
<td>Ile</td>
<td>Cys</td>
<td>Glu</td>
<td>Leu</td>
<td>Ile</td>
<td>Gln</td>
<td>Ala</td>
<td>Phe</td>
<td>Glu</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Asp</td>
<td>Phe</td>
<td>Gln</td>
<td>Glu</td>
<td>Ser</td>
<td>Ala</td>
<td>Asp</td>
<td>Ser</td>
<td>Phe</td>
<td>Leu</td>
<td>Met</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>His</td>
<td>Ala</td>
<td>Tyr</td>
<td>Gln</td>
<td>Gly</td>
<td>Asp</td>
<td>Tyr</td>
<td>Lys</td>
<td>Gln</td>
<td>Phe</td>
<td>Leu</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Lys</td>
<td>Tyr</td>
<td>Leu</td>
<td>Glu</td>
<td>Gly</td>
<td>His</td>
<td>Gly</td>
<td>Phe</td>
<td>Arg</td>
<td>Phe</td>
<td>Glu</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Gly</td>
<td>Val</td>
<td>Lys</td>
<td>Arg</td>
<td>Leu</td>
<td>Glu</td>
<td>Glu</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
<td>Ala</td>
</tr>
</tbody>
</table>
-continued

Pro Lys Lys Asp Ser Thr Gln Asn Asn Asp Asn Pro Ala Gln Arg Ala 530 535 540
Gln Glu Tyr Ala Arg Asp Asn Ile Gln Asp Thr Pro Thr Pro His Arg 545 550 555 560
Ala Leu Thr Pro Ile Ser Glu Thr Gly Ser Asn Gly His Asn Glu 565 570 575
Asp Asp Ile Asp Ser Ile Pro Pro Leu Glu Ser Asp Glu Gln Asn Asn 580 585 590
Thr Glu Thr Thr Ile Thr Thr Lys Asn Thr Thr Ala Pro Pro Ala 595 600 605
Pro Val Tyr Arg Ser Asn Ser Glu Gly Pro Leu Pro Gln Glu Lys 610 615 620
Ser Gln Lys Gln Pro Asn Gln Val Ser Gly Ser Glu Asn Thr Asp Asn 625 630 635 640
Lys Pro His Ser Glu Gln Ser Val Glu Met Tyr Arg His Ile Leu 645 650 655
Gln Thr Gln Gly Pro Phe Asp Ala Ile Leu Tyr Tyr Tyr Met Met Thr 660 665 670
Glu Glu Pro Ile Val Phe Ser Thr Ser Asp Gly Lys Glu Tyr Val Tyr 675 680 685
Pro Asp Ser Leu Glu Gly Glu His Pro Pro Trp Leu Ser Glu Gln Lys Glu 690 695 700
Ala Leu Asn Glu Asp Asn Arg Phe Ile Thr Met Asp Asp Gln Gln Phe 705 710 715 720
Tyr Trp Pro Val Met Asn His Arg Asn Lys Phe Met Ala Ile Leu Glu 725 730 735
His Lys

<210> SEQ ID NO 12
<211> LENGTH: 341
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc feature
<223> OTHER INFORMATION: Cote d’Ivoire ebolavirus VP35 NP protein

<400> SEQUENCE: 12

Met Ile Ser Thr Arg Ala Ala Ala Ile Asn Asp Pro Ser Leu Pro Ile 1 5 10 15
Arg Asn Gln Cys Thr Arg Gly Pro Glu Leu Ser Gly Trp Ile Ser Glu 20 25 30
Gln Leu Met Thr Gly Lys Ile Pro Val His Glu Ile Phe Asn Asp Thr 35 40 45
Glu Pro His Ile Ser Ser Gly Ser Asp Cys Leu Pro Arg Pro Lys Asn 50 55 60
Thr Ala Pro Arg Thr Arg Asn Thr Gln Thr Glu Thr Asp Pro Val Cys 65 70 75 80
Asn His Asn Phe Glu Asp Val Thr Gln Ala Leu Thr Ser Leu Thr Asn 95 90 95
Val Ile Gln Lys Gln Ala Leu Asn Leu Glu Ser Leu Glu Gln Arg Ile 100 105 110
Ile Asp Leu Glu Asn Gly Leu Lys Pro Met Tyr Asp Met Ala Lys Val 115 120 125
Ile Ser Ala Leu Asn Arg Ser Cys Ala Glu Met Val Ala Lys Tyr Asp
130 135 140
Leu Leu Val Met Thr Thr Gly Arg Ala Thr Ala Ala Ala Ala Thr
145 150 155 160
Glu Ala Tyr Trp Glu Glu His Gly Gln Pro Pro Pro Gly Pro Ser Leu
165 170 175
Tyr Glu Glu Ser Ala Ile Arg Gly Lys Ile Asn Lys Glu Glu Asp Lys
180 185 190
Val Pro Lys Glu Val Gln Ala Phe Arg Asn Leu Asp Ser Thr Ser
195 200 205
Ser Leu Thr Glu Asn Phe Gly Lys Pro Asp Ile Ser Ala Lys Asp
210 215 220
Leu Arg Asp Ile Met Tyr Asp His Leu Pro Gly Phe Gly Thr Ala Phe
225 230 235 240
His Gln Leu Val Gln Val Ile Lys Leu Gly Lys Asp Asn Ser Ala
245 250 255
Leu Asp Ile Ile His Ala Glu Phe Glu Ala Ser Leu Ala Glu Gly Asp
260 265 270
Ser Pro Gln Cys Ala Leu Ile Gln Ile Thr Lys Arg Ile Pro Ile Phe
275 280 285
Gln Asp Ala Thr Pro Pro Thr Ile His Ile Arg Ser Arg Gly Asp Ile
290 295 300
Pro Arg Ala Cys Gln Lys Ser Leu Arg Pro Val Pro Pro Ser Pro Lys
305 310 315 320
Ile Asp Arg Gly Trp Val Cys Ile Phe Glu Leu Gln Asp Gly Lys Thr
325 330 335
Leu Gly Leu Lys Ile
340
<210> SEQ ID NO 13
<211> LENGTH: 326
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE: misc_feature
<223> OTHER INFORMATION: Cote d'Ivoire ebolavirus VP40 NP protein
<400> SEQUENCE: 13
Met Arg Arg Ile Ile Leu Pro Thr Ala Pro Pro Glu Tyr Met Glu Ala
1  5 10 15
Val Tyr Pro Met Arg Thr Met Asn Ser Gly Ala Asp Asn Thr Ala Ser
20 25 30
Gly Pro Asn Tyr Thr Thr Gly Val Met Thr Asn Thr Pro Ser
35 40 45
Asn Ser Leu Arg Pro Val Ala Asp Asn Ile Asp His Pro Ser His
50 55 60
Thr Pro Asn Ser Val Ala Ser Ala Phe Ile Leu Glu Ala Met Val Asn
65 70 75 80
Val Ile Ser Gly Pro Lys Val Leu Met Lys Glu Ile Pro Ile Trp Leu
85 90 95
Pro Leu Gly Val Ser Asp Glu Thr Tyr Ser Phe Asp Ser Thr Thr
100 105 110
Ala Ala Ile Met Leu Ala Ser Tyr Thr Ile Thr His Phe Gly Lys Thr
115 120...
Ser Asn Pro Leu Val Arg Ile Asn Arg Leu Gly Pro Gly Ile Pro Asp
115 120 125
His Pro Leu Arg Leu Arg Ile Gly Asn Gin Ala Phe Leu Gin Glu
130 135 140
Phe Val Leu Pro Pro Val Gin Leu Pro Gin Tyr Phe Thr Phe Asp Leu
145 150 155 160
Thr Ala Leu Lys Leu Ile Thr Gin Pro Leu Pro Ala Ala Thr Thr
165 170 175
Asp Glu Thr Pro Ala Val Ser Thr Gly Thr Leu Arg Pro Gly Ile Ser
180 185 190
Phe His Pro Lys Leu Arg Pro Ile Leu Pro Gly Arg Ala Gly Lys
195 200 205
Lys Gly Ser Asn Ser Asp Leu Thr Ser Pro Asp Lys Ile Gin Ala Ile
210 215 220
Met Asn Phe Leu Gin Asp Leu Ile Val Pro Ile Asp Pro Thr Lys
225 230 235 240
Asn Ile Met Gly Ile Glu Val Pro Glu Leu Leu Val His Arg Leu Thr
245 250 255
Gly Lys Thr Thr Thr Lys Asn Gly Gin Pro Ile Ile Pro Ile Leu
260 265 270
Leu Pro Lys Tyr Ile Gly Leu Asp Pro Leu Ser Gin Gly Asp Leu Thr
275 280 285
Met Val Ile Thr Gin Asp Cys Asp Ser Cys His Ser Pro Ala Ser Leu
290 295 300
Pro Pro Val Asn Glu Lys
305 310 315 320

<210> SEQ ID NO 14
<211> LENGTH: 676
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> OTHER INFORMATION: Cote d'Ivoire ebolavirus GP NP protein
<400> SEQUENCE: 14
Met Gly Ala Ser Gly Ile Leu Gin Leu Pro Arg Glu Arg Phe Arg Lys
1  5  10  15
Thr Ser Phe Phe Val Trp Val Ile Leu Phe His Lys Val Phe Ser
20 25 30
Ile Pro Leu Gly Val Val His Asn Thr Leu Gin Val Ser Asp Ile
35 40 45
Asp Lys Phe Val Cys Arg Asp Lys Leu Ser Ser Thr Ser Gin Leu Lys
50 55 60
Ser Val Gly Leu Asn Leu Gin Gly Gin Val Ala Thr Gin Leu Pro
65 70 75 80
Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gin Val Pro Pro Lys Val
85 90 95
Val Asn Cys Glu Ala Gin Glu Trp Ala Gin Asn Cys Tyr Asn Ile Leu Ala
100 105 110
Ile Lys Lys Val Asp Gly Ser Glu Gin Ile Gin Gin Gin Gin Gin Gin Gin Glu
115 120 125
-continued

Val Arg Asp Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
110 135
Gly Pro Cys Pro Gly Gly Leu Ala Phe His Lys Glu Gly Ala Phe Phe
145 150 155 160
Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Gly Thr Thr Phe
165 170 175
Ala Glu Gly Val Ile Ala Phe Leu Ile Leu Pro Lys Ala Arg Lys Asp
180 185 190
Phe Phe Gln Ser Pro Pro Leu His Glu Pro Ala Asn Met Thr Thr Asp
195 200 205
Pro Ser Ser Tyr Tyr His Thr Thr Ile Asn Tyr Val Asp Asn
210 215 220
Phe Gly Thr Asn Thr Glu Phe Leu Phe Gln Val Asp His Leu Thr
225 230 235 240
Tyr Val Gln Leu Glu Ala Arg Phe Thr Pro Gln Phe Leu Val Leu Leu
245 250 255
Asn Glu Thr Ile Tyr Ser Asp Asn Arg Arg Ser Asn Thr Thr Gly Lys
260 265 270
Leu Ile Trp Lys Ile Asn Pro Thr Val Asp Thr Ser Met Gly Glu Trp
275 280 285
Ala Phe Trp Glu Asn Lys Asn Phe Thr Lys Leu Ser Ser Glu
290 295 300
Glu Leu Ser Phe Val Pro Val Pro Glu Thr Gin Asn Gin Val Leu Asp
305 310 315 320
Thr Thr Ala Thr Val Ser Pro Pro Ile Ser Ala His Asn His Ala Ala
325 330 335
Glu Asp His Lys Glu Leu Val Ser Glu Asp Ser Thr Pro Val Val Gin
340 345 350
Met Gin Asn Ile Lys Gly Lys Asp Thr Met Pro Thr Thr Val Thr Gly
360 365
Val Pro Thr Thr Thr Pro Ser Pro Phe Pro Ile Asn Ala Arg Asn Thr
370 375 380
Asp His Thr Lys Ser Phe Ile Gly Glu Gly Pro Gin Glu Asp His
385 390 395 400
Ser Thr Thr Gin Pro Ala Lys Thr Thr Gin Pro Thr Asn Ser Thr
405 410 415
Glu Ser Thr Thr Leu Asn Pro Thr Ser Glu Pro Ser Ser Arg Gly Thr
420 425 430
Gly Pro Ser Ser Pro Thr Val Pro Leu Thr Ser Gin Ser Ser Gin Thr
435 440 445
Leu Gly Lys Thr Thr Pro Thr Thr Leu Pro Glu Gin His Thr Ala Ala
450 455 460
Ser Ala Ile Pro Arg Ala Val His Pro Asp Glu Leu Ser Gly Pro Gly
465 470 475 480
Phe Leu Thr Asn Thr Ile Arg Gly Val Thr Asn Leu Thr Gly Ser
485 490 495
Arg Arg Lys Arg Arg Asp Val Thr Pro Asn Thr Gin Pro Lys Cys Asn
500 505 510
Pro Asn Leu His Tyr Trp Thr Ala Leu Asp Glu Gly Ala Ala Ile Gly
515 520 525
Leu Ala Trp Ile Pro Tyr Phe Gly Pro Ala Ala Glu Gly Ile Tyr Thr
Glu Gly Ile Met Glu Asn Glu Asn Leu Ile Cys Gly Leu Arg Gln
545 550 555 560
Leu Ala Asn Glu Thr Thr Glu Ala Leu Glu Leu Phe Leu Arg Ala Thr
565 570 575
Thr Glu Leu Arg Thr Phe Ser Ile Leu Asn Arg Lys Ala Ile Asp Phe
590
Leu Leu Gln Arg Trp Gly Gly Thr Cys His Ile Leu Gln Pro Asp Cys
595 600 605
Cys Ile Glu Pro Gln Asp Trp Thr Lys Asn Ile Thr Asp Lys Ile Asp
610 615 620
Gln Ile Ile His Asp Phe Val Asp Asn Asn Leu Pro Asn Gln Asn Asp
625 630 635 640
Gly Ser Asn Trp Trp Thr Gly Trp Lys Gln Trp Val Pro Ala Gly Ile
645 650 655
Gly Ile Thr Gly Val Ile Ile Ala Ile Ile Ala Leu Leu Cys Ile Cys
660 665 670
Lys Phe Met Leu
675
–continued

<210> SEQ ID NO 15
<211> LENGTH: 365
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc.feature
<222> OTHER INFORMATION: Cote d’Ivoire ebolavirus GPF NP protein
<400> SEQUENCE: 15
Met Gly Ala Ser Gly Ile Leu Gln Leu Pro Arg Glu Arg Phe Arg Lys
1  5  10  15
Thr Ser Phe Phe Val Trp Val Ile Leu Phe His Lys Val Phe Ser
20 25 30
Ile Pro Leu Gly Val His Asn Thr Leu Gln Val Ser Asp Ile
35 40 45
Asp Lys Phe Val Cys Arg Asp Lys Ser Ser Ser Glu Gln Leu Lys
50 55 60
Ser Val Gly Leu Asn Leu Gln Gly Asn Gly Val Ala Thr Asp Val Pro
65 70 75 80
Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val
85 90 95
Val Asn Cys Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Ala
100 105 110
Ile Lys Lys Val Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly
115 120 125
Val Arg Asp Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
130 135 140
Gly Pro Cys Pro Gly Gly Leu Ala Phe His Lys Glu Gly Ala Phe Phe
145 150 155 160
Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Gly Thr Thr Phe
165 170 175
Ala Glu Gly Val Ile Ala Phe Leu Ile Leu Pro Lys Ala Arg Lys Asp
180 185 190
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met Gly Ala Ser Gly Ile Leu Gln Leu Pro Arg Glu Arg Phe Arg Lys</td>
<td>1 5 10 15</td>
</tr>
<tr>
<td>Thr Ser Phe Phe Val Trp Val Ile Leu Phe His Lys Val Phe Ser</td>
<td>20 25 30</td>
</tr>
<tr>
<td>Ile Pro Leu Gly Val His Asn Thr Leu Gln Val Ser Asp Ile</td>
<td>35 40 45</td>
</tr>
<tr>
<td>Asp Lys Phe Val Cys Arg Asp Lys Leu Ser Ser Thr Ser Gin Leu Lys</td>
<td>50 55 60</td>
</tr>
<tr>
<td>Ser Val Gly Leu Asn Leu Glu Gly Asn Gin Val Ala Thr Asp Val Pro</td>
<td>65 70 75 80</td>
</tr>
<tr>
<td>Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val</td>
<td>85 90 95</td>
</tr>
<tr>
<td>Val Asn Cys Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Ala</td>
<td>100 105 110</td>
</tr>
<tr>
<td>Ile Lys Lys Val Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly</td>
<td>115 120 125</td>
</tr>
<tr>
<td>Val Arg Asp Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr</td>
<td>130 135 140</td>
</tr>
<tr>
<td>Gly Pro Cys Pro Gly Gly Leu Ala Phe His Lys Glu Gly Ala Phe</td>
<td>145 150 155 160</td>
</tr>
<tr>
<td>Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Gly Thr Phe</td>
<td>165 170 175</td>
</tr>
<tr>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO 17**
**<211> LENGTH: 289**
**<212> TYPE: PRT**
**<213> ORGANISM: Bundibugyo ebolavirus**
**<220> FEATURE: misc_feature**
**<221> OTHER INFORMATION: Cote d'Ivoire ebolavirus VP30 NP protein**

**<400> SEQUENCE: 17**

<table>
<thead>
<tr>
<th>Met</th>
<th>Glu</th>
<th>Val</th>
<th>Val</th>
<th>His</th>
<th>Glu</th>
<th>Arg</th>
<th>Ser</th>
<th>Arg</th>
<th>Ser</th>
<th>Arg</th>
<th>Ser</th>
<th>Arg</th>
<th>Gln</th>
<th>Asn</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Thr</td>
<td>Arg</td>
<td>Asp</td>
<td>Gly</td>
<td>Pro</td>
<td>Ser</td>
<td>His</td>
<td>Leu</td>
<td>Val</td>
<td>Arg</td>
<td>Ala</td>
<td>Arg</td>
<td>Ser</td>
<td>Ser</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td>Tyr</td>
<td>Arg</td>
<td>Ser</td>
<td>Glu</td>
<td>Tyr</td>
<td>Thr</td>
<td>Pro</td>
<td>Arg</td>
<td>Ser</td>
<td>Ala</td>
<td>Ser</td>
<td>Gln</td>
<td>Ile</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Val</td>
<td>Pro</td>
<td>Thr</td>
<td>Val</td>
<td>Phe</td>
<td>His</td>
<td>Arg</td>
<td>Lys</td>
<td>Thr</td>
<td>Asp</td>
<td>Leu</td>
<td>Leu</td>
<td>Thr</td>
<td>Val</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td>Ala</td>
<td>Pro</td>
<td>Lys</td>
<td>Cys</td>
<td>Pro</td>
<td>Thr</td>
<td>Leu</td>
<td>Lys</td>
<td>Gly</td>
<td>Phe</td>
<td>Leu</td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Asp</td>
<td>Ser</td>
<td>Asn</td>
<td>Phe</td>
<td>Cys</td>
<td>Lys</td>
<td>Lys</td>
<td>Asp</td>
<td>His</td>
<td>Gln</td>
<td>Leu</td>
<td>Glu</td>
<td>Ser</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Leu</td>
<td>Ala</td>
<td>Arg</td>
<td>Lys</td>
<td>Thr</td>
<td>Cys</td>
<td>Gly</td>
<td>Ser</td>
<td>Thr</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Gln</td>
<td>Gln</td>
<td>Leu</td>
<td>Ser</td>
<td>Ile</td>
<td>Val</td>
<td>Ala</td>
<td>Pro</td>
<td>Lys</td>
<td>Asp</td>
<td>Ser</td>
<td>Arg</td>
<td>Ala</td>
<td>Asn</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ile</td>
<td>Ala</td>
<td>Glu</td>
<td>Asp</td>
<td>Phe</td>
<td>Gln</td>
<td>Gln</td>
<td>Lys</td>
<td>Asp</td>
<td>Gln</td>
<td>Pro</td>
<td>Lys</td>
<td>Val</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Met</td>
<td>Leu</td>
<td>Ile</td>
<td>Glu</td>
<td>Thr</td>
<td>Ala</td>
<td>Glu</td>
<td>Tyr</td>
<td>Trp</td>
<td>Ser</td>
<td>Lys</td>
<td>Gln</td>
<td>Asp</td>
<td>Ile</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Thr</td>
<td>Arg</td>
<td>Lys</td>
<td>Phe</td>
<td>Ser</td>
<td>Lys</td>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
<td>Ser</td>
<td>Leu</td>
<td>Leu</td>
<td>Cys</td>
<td>Glu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Leu</td>
<td>Arg</td>
<td>Arg</td>
<td>Glu</td>
<td>Gly</td>
<td>Leu</td>
<td>Gly</td>
<td>Phe</td>
<td>Asp</td>
<td>Gin</td>
<td>Ser</td>
<td>Glu</td>
<td>Ser</td>
<td>Val</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Val</td>
<td>Tyr</td>
<td>Glu</td>
<td>Arg</td>
<td>Leu</td>
<td>His</td>
<td>Ser</td>
<td>Asp</td>
<td>Lys</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Phe</td>
<td>Glu</td>
<td>Ala</td>
</tr>
</tbody>
</table>
-continued

Ala Leu Trp Gln Gln Trp Asp Arg Gln Ser Leu Ile Met Phe Ile Thr 225 230 235 240
Ala Phe Leu Arg Ile Ala Leu Glu Leu Pro Cys Glu Ser Ser Ser Val 245 250 255
Val Ile Ser Gly Leu Arg Met Leu Ile Pro Glu Ser Glu Ala Thr Glu 260 265 270
Val Val Thr Pro Ser Glu Thr Cys Thr Trp Ser Glu Gly Gly Ser Ser 275 280 285
His

SEQ ID NO: 18
LENGTH: 251
TYPE: PRT
ORGANISM: Bundibugyo ebolavirus
FEATURE: misc_feature
OTHER INFORMATION: Cote d'Ivoire ebolavirus VP44 NP protein

SEQUENCE: 18

Met Ala Lys Ala Thr Gly Arg Tyr Asn Leu Ile Ser Pro Lys Lys Asp 1 5 10 15
Leu Glu Lys Gly Leu Val Leu Asp Leu Cys Thr Leu Ser Val Ala 20 25 30
Gln Thr Val Gln Gly Trp Lys Val Thr Trp Ala Gly Ile Glu Phe Asp 35 40 45
Val Thr Gln Lys Gly Met Ala Leu Leu His Arg Leu Lys Thr Ser Asp 50 55 60
Phe Ala Pro Ala Trp Ser Met Thr Arg Asn Leu Phe Pro His Leu Phe 65 70 75 80
Gln Asn Pro Asn Ser Thr Ile Glu Ser Pro Leu Trp Ala Leu Arg Val 85 90 95
Ile Leu Ala Ala Glu Ile Glu Asp Glu Leu Ile Asp Glu Ser Leu Ile 100 105 110
Glu Pro Leu Ala Gly Ala Leu Gly Leu Ile Ala Asp Trp Leu Leu Thr 115 120 125
Thr Gly Thr Asn His Phe Glu Met Arg Thr Gln Glu Ala Lys Glu Gln 130 135 140
Leu Ser Leu Lys Met Leu Ser Ser Asn Ile Leu Lys Phe 145 150 155 160
Ile Asn Glu Leu Asp Ala His Val Asn Tyr Asn Gly Leu Leu 165 170 175
Ser Ser Ile Glu Ile Gly Thr Lys Ser His Thr Ile Ile Thr Arg 180 185 190
Thr Asn Met Gly Phe Leu Val Glu Leu Glu Glu Pro Asp Lys Ser Ala 195 200 205
Met Asn Thr Arg Lys Pro Gly Pro Val Lys Phe Ser Leu Leu His Glu 210 215 220
Ser Thr Leu Lys Thr Leu Ala Lys Lys Pro Ala Thr Glu Met Glu Ala 225 230 235 240
Leu Ile Leu Glu Phe Asn Ser Ser Leu Ala Ile 245 250
<210> SEQ ID NO: 19
<211> LENGTH: 2210
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE: misc_feature
<221> NAME/KEY:
<222> OTHER INFORMATION: Cote d’Ivoire ebolavirus NP protein

<400> SEQUENCE: 19

Met Ala Thr Gln His Thr Gln Tyr Pro Asp Ala Arg Leu Ser Ser Pro
1  5  10  15
Ile Val Leu Asp Gln Cys Asp Leu Val Thr Arg Ala Cys Gly Leu Tyr
20  25  30
Ser Ala Tyr Ser Leu Asn Pro Gln Leu Lys Asn Cys Arg Leu Pro Lys
35  40  45
His Ile Tyr Arg Leu Lys Tyr Asp Thr Thr Val Thr Glu Phe Leu Ser
50  55  60
Asp Val Pro Val Ala Thr Leu Pro Ala Asp Phe Leu Val Pro Thr Phe
65  70  75  80
Leu Arg Thr Leu Ser Gly Asn Gly Ser Cys Pro Ile Asp Pro Lys Cys
85  90  95
Ser Gln Phe Leu Glu Glu Val Asn Tyr Thr Leu Gln Asp Ile Arg
100 105 110
Phe Leu Asn Tyr Tyr Leu Asn Arg Ala Gly Val His Asn Asp His Val
115 120 125
Asp Arg Asp Phe Gly Gln Lys Ile Arg Asn Leu Ile Cys Asp Asn Glu
130 135 140
Val Leu His Gln Met Phe His Trp Tyr Asp Ala Ile Leu Ala Arg
145 150 155 160
Arg Gly Arg Leu Asn Arg Gly Asn Arg Ser Thr Trp Phe Ala Ser
165 170 175
Asp Asn Leu Val Asp Ile Leu Gly Tyr Gly Asp Tyr Ile Phe Trp Lys
180 185 190
Ile Pro Leu Ser Leu Leu Pro Val Asp Thr Gln Gly Leu Pro His Ala
195 200 205
Ala Lys Asp Thr Tyr His Glu Ser Val Phe Lys Glu Ala Ile Gln Gly
210 215 220
His Thr His Ile Val Ser Ile Ser Thr Ala Asp Val Leu Ile Met Cys
225 230 235 240
Lys Asp Ile Ile Thr Cys Arg Phe Asn Thr Leu Leu Ile Ala Ala Val
245 250 255
Ala Asn Leu Glu Asp Ser Val His Ser Asp Tyr Pro Leu Pro Glu Thr
260 265 270
Val Ser Asp Leu Tyr Lys Ala Gly Asp Tyr Leu Ile Ser Leu Leu Gly
275 280 285
Ser Glu Gly Tyr Lys Val Ile Lys Phe Leu Glu Pro Leu Cys Leu Ala
290 295 300
Lys Ile Gln Leu Cys Ser Asn Tyr Thr Glu Arg Lys Gly Arg Phe Leu
305 310 315 320
Thr Gln Met His Leu Ala Val Asn His Thr Leu Glu Leu Thr Gly
325 330 335
Ser Arg Glu Leu Arg Pro Gln Gln Ile Arg Lys Val Arg Glu Phe His
340 345 350
-continued

Gln Met Leu Ile Asn Leu Lys Ala Thr Pro Gin Gln Leu Cys Glu Leu 355 360 365
Phe Ser Val Gin Lys His Trp Gly His Pro Val Leu His Ser Glu Lys 370 375 380
Ala Ile Gln Lys Val Lys His Ala Thr Val Ile Lys Ala Leu Arg 385 390 395 400
Pro Ile Ile Ile Phe Glu Thr Tyr Cys Val Phe Lys Tyr Ser Ile Ala 405 410 415
Lys His Tyr Phe Asp Ser Gin Gly Thr Trp Tyr Ser Val Thr Ser Asp 420 425 430
Arg Cys Leu Thr Pro Gly Leu Ser Ser Tyr Ile Lys Arg Asn Gin Phe 435 440 445
Pro Pro Leu Pro Met Ile Lys Glu Leu Trp Glu Phe Tyr His Leu 450 455 460
Asp His Pro Pro Leu Phe Ser Thr Lys Val Ile Ser Asp Leu Ser Ile 465 470 475 480
Phe Ile Lys Asp Arg Ala Thr Ala Val Glu Lys Thr Cys Trp Asp Ala 485 490 495
Val Phe Glu Pro Asn Val Leu Gly Tyr Asn Pro Pro Asn Lys Phe Ala 500 505 510
Thr Lys Arg Val Pro Glu Gin Phe Leu Gin Gin Glu Asn Phe Ser Ile 515 520 525
Glu Ser Val Leu His Tyr Ala Gin Leu Glu Tyr Leu Pro Glu 530 535 540
Tyr Arg Asn Phe Ser Phe Ser Leu Lys Glu Leu Asn Ile Gly 545 550 555 560
Arg Ala Phe Gly Lys Leu Pro Tyr Pro Thr Arg Asn Val Gin Thr Leu 565 570 575
Cys Glu Ala Leu Leu Ala Asp Gly Leu Ala Lys Ala Phe Pro Ser Asn 580 585 590
Met Met Val Val Thr Glu Arg Glu Gin Lys Glu Ser Leu Leu His Gin 595 600 605
Ala Ser Trp His His Thr Ser Asp Asp Phe Gly Glu Asn Ala Thr Val 610 615 620
Arg Gly Ser Ser Phe Val Thr Asp Leu Gly Tyr Asn Leu Ala Phe 625 630 635 640
Arg Tyr Glu Phe Thr Ala Pro Phe Ile Glu Tyr Cys Asn Arg Cys Tyr 645 650 655 660
Gly Val Arg Asn Leu Phe Asn Trp Met His Tyr Thr Ile Pro Gin Cys 665 670
Tyr Ile His Val Ser Asp Tyr Asn Pro Pro His Gly Val Ser Leu 675 680 685
Glu Asn Arg Glu Asn Pro Pro Glu Gly Pro Ser Ser Tyr Arg Gly His 690 695 700
Leu Gly Lys Glu Leu Gin Gin Lys Leu Trp Thr Ser Ile Ser 705 710 715 720
Cys Ala Gln Ile Ser Leu Val Glu Ile Lys Thr Gly Phe Lys Leu Arg 725 730 735 740
Ser Ala Val Met Gly Asp Asn Gin Cys Ile Thr Val Leu Ser Val Phe 745 750
Pro Leu Glu Thr Glu Ser Ser Glu Gin Glu Leu Ser Ser Glu Asp Asn
Ala Ala Arg Val Ala Ala Ser Leu Ala Lys Val Thr Ser Ala Cys Gly
770 775 780
Ile Phe Leu Lys Pro Asp Glu Thr Phe Val His Ser Gly Phe Ile Tyr
785 790 795 800
Phe Gly Lys Lys Gln Tyr Leu Asn Gly Val Gin Leu Pro Gin Ser Leu
805 810 815
Lys Thr Ala Thr Arg Ile Ala Pro Leu Ser Asp Ala Ile Phe Asp Asp
820 825 830
Leu Gin Gly Thr Leu Ala Ser Ile Gly Thr Ala Phe Glu Arg Ser Ile
835 840 845
Ser Gin Thr Arg His Val Val Pro Cyg Arg Val Ala Ala Ala Phe His
850 855 860
Thr Phe Phe Ser Val Arg Ile Leu Gin Tyr Asp To His Leu Gly Phe Asn
865 870 875 880
Lys Gin Thr Leu Gin Leu Ser Leu Ser Lys Pro Leu Asp Phe
885 890 895
Gly Thr Ile Thr Leu Ala Leu Ala Val Pro Gin Val Leu Gly Gly Leu
900 905 910
Ser Gin Thr Leu Asn Pro Gin Lys Cys Phe Tyr Gin Asn Leu Gly Gin Asp Pro
915 920 925
Val Thr Ser Gin Leu Lys Thr Tyr Leu Gin Met Ile His
930 935 940
Met Asp Gin Phe Leu Pro Leu Ile Ala Lys Gin Gin Pro Gin Cys Gin
945 950 955 960
Ser Ala Ile Gin Phe Val Leu Gin Pro Ser Gin Leu Asn Val Pro Gin
965 970 975
Ser Gin Asp Leu Thr Ser Phe Leu Gin Gin Ile Val Arg Arg Thr Ile
980 985 990
Thr Leu Ser Ala Gin Gin Leu Gin Asn Thr Leu Gin Gin Ser Ser
995 1000 1005
Ala Asp Leu Gin Asp Gin Gin Val Cyg Lys Trp Leu Leu Ser Gin
1010 1015 1020
Thr Pro Val Met Ser Arg Phe Ala Ala Gin Ile Phe Ser Arg Thr
1025 1030 1035
Pro Ser Gin Lys Arg Leu Gin Ile Leu Gin Tyr Leu Glu Gly Thr
1040 1045 1050
Arg Thr Leu Leu Ala Gin Gin His Gin Met Gin Thr Gin Thr
1055 1060 1065
Pro Ile Gin Asp Gin Gin Leu Gin Gin Ile Thr Leu Leu Gin Gin Gin Gin
1070 1075 1080
Leu Trp Phe Ser Tyr Leu Asp His Cys Gin Gin Val Gin Leu Gin
1085 1090 1095
Ala Leu Thr Gin Ile Thr Cys Thr Val Gin Gin Gin Gin Gin
1100 1105 1110
Arg Gin Tyr Thr Trp Gin Gin Gin Gin Gin Gin Gin Gin
1115 1120 1125
Gly Gin Thr Gin Gin Gin Gin Gin Gin Gin Gin
1130 1135 1140
Leu Lys Pro Gin Gin His Cys Pro Gin Cys Gin Gin Gin Gin Gin
1145 1150 1155
| 1536 | Ile Gly Gly Thr Ala Gly Asp Arg Gly Leu Ser Asp Ala Ala Arg  |
| 1540 |                           |                            |
| 1544 |                           |                            |
| 1590 | Leu Phe Leu Arg Thr Ala Ile Thr Val Phe Leu Gln Phe Val Arg  |
| 1594 |                           |                            |
| 1598 |                           |                            |
| 1602 | Lys Trp Ile Val Glu Arg Lys Thr Ala Ile Pro Leu Trp Val Ile  |
| 1606 |                           |                            |
| 1610 |                           |                            |
| 1614 |                           |                            |
| 1618 |                           |                            |
| 1622 | His Val Ile Ala Leu Leu Gln His Glu Ser Ser His Asp His      |
| 1626 |                           |                            |
| 1630 | Val Cys Ala Ala Glu Ala His Ser Arg Val Glu Thr Phe Asp Asn  |
| 1634 |                           |                            |
| 1638 |                           |                            |
| 1642 | Leu Val Tyr Met Cys Lys Ser Thr Ala Ser Asn Phe Phe His Ala  |
| 1646 |                           |                            |
| 1650 |                           |                            |
| 1654 | Ser Leu Ala Tyr Trp Arg Ser Arg Ser Lys Asn Gln Asp Lys Arg  |
| 1658 |                           |                            |
| 1662 |                           |                            |
| 1666 | Glu Met Thr Lys Ile Leu Ser Leu Thr Gln Thr Glu Lys Lys Asn  |
| 1670 |                           |                            |
| 1674 |                           |                            |
| 1678 |                           |                            |
| 1682 | Ser Phe Gly Tyr Thr Ala His Pro Glu Ser Thr Ala Val Leu Gly  |
| 1686 |                           |                            |
| 1690 |                           |                            |
| 1694 | Ser Leu Gln Thr Ser Leu Ala Pro Pro Pro Ser Ala Asp Glu Ala  |
| 1698 |                           |                            |
| 1702 |                           |                            |
| 1706 | Thr Tyr Asp Arg Lys Asn Lys Val Leu Lys Ala Ser Arg Pro Gly  |
| 1710 |                           |                            |
| 1714 |                           |                            |
| 1718 | Lys Tyr Ser Gln Asn Thr Thr Lys Ala Pro Pro Asn Gln Thr Ser  |
| 1722 |                           |                            |
| 1726 |                           |                            |
| 1730 | Cys Arg Asp Val Ser Pro Asn Ile Thr Gly Thr Asp Gly Cys Pro  |
| 1734 |                           |                            |
| 1738 |                           |                            |
| 1742 | Ser Ala Asn Glu Gly Ser Asn Ser Asn Asn Asn Leu Val Ser      |
| 1746 |                           |                            |
| 1750 |                           |                            |
| 1754 | His Arg Ile Val Leu Pro Phe Phe Thr Leu Ser His Asn Tyr Asn  |
| 1758 |                           |                            |
| 1762 |                           |                            |
| 1766 | Glu Arg Pro Ser Ile Arg Lys Ser Gly Thr Thr Glu Ile Val      |
| 1770 |                           |                            |
| 1774 |                           |                            |
| 1778 | Arg Leu Thr Arg Gln Leu Arg Ala Ile Pro Asp Thr Thr Ile Tyr  |
| 1782 |                           |                            |
| 1786 |                           |                            |
| 1790 | Cys Arg Phe Thr Gly Ile Val Ser Ser Met His Tyr Lys Leu Asp  |
| 1794 |                           |                            |
| 1800 |                           |                            |
| 1804 | Glu Val Leu Trp Glu Phe Asp Asn Phe Lys Ser Ala Ile Thr Leu  |
| 1808 |                           |                            |
| 1812 |                           |                            |
| 1816 | Ala Glu Gly Glu Gly Ser Gly Ala Leu Leu Leu Leu Gln Lys Tyr  |
| 1820 |                           |                            |
| 1824 |                           |                            |
| 1828 | Lys Val Glu Thr Leu Phe Phe Asn Thr Leu Ala Thr Glu His Ser  |
| 1832 |                           |                            |
| 1836 |                           |                            |
| 1840 | Ile Glu Ala Glu Ile Ser Gly Ile Thr Thr Pro Arg Met Leu      |
| 1844 |                           |                            |
| 1848 |                           |                            |
| 1852 | Leu Pro Ile Met Ser Arg Phe His Gly Gly Gln Ile Lys Val Thr  |
| 1856 |                           |                            |
| 1860 |                           |                            |
| 1864 | Leu Asn Asn Ser Ala Ser Gln Ile Thr Asp Ile Thr Asn Pro Ser  |
| 1868 |                           |                            |
| 1872 |                           |                            |
| 1876 | Trp Leu Ala Asp Gln Lys Ser Arg Ile Pro Lys Gln Val Glu Ile  |
<210> SEQ ID NO 20
<211> LENGTH: 18959
<212> TYPE: DNA
<213> ORGANISM: Zaire ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: Full viral sequence
<400> SEQUENCE: 20

ctgagcaca  aaaaagaga  agaagttttta  ggtttttttttttt  tgtttctgata  actatggs
agatattaa  ttttctcttc  attgaaatatt  atacggsat  ttaaatggaa  atgttactg
taataccgacc  tgggttgttt  cacagcaca  tacaaagat  agagaacac  ctaggtctcc  180
gagaagccca  aggccatcag  tgtgtcagt  tgaatactcc  tgtcaacac  ctaggttta  240
tcaatcaca  agttcctctc  cagacccctg  aggggtaccc  aatcacttta  atagaaaaat  300
tatggctaaa  ggcagcagct  agttcagct  caaaaaagcc  aagttggaga  ttaaccttgg  360
ttttgaacct  gacacctgtg  ggattgaag  attcacaac  octaaagct  gggttaaa  420
attggaaata  gttaaagac  aatcgttcg  gaatacaca  attcgagta  tgtattctcg  480
tcctcaagaa  atcttggtgg  gcgcaggtct  cacagatctt  gacaaggtt  acaacaagat  540
cctgagacgc  gcgtcttccg  ttcacaaaggg  gattgttcgg  caagaggtca  tccaccagta  600
tcaagtaaa  aaccttggaag  aataattgcca  acctactata  cagggccttg  aagoaagttg  660
tgattaacaa  gagagtggcg  acagttttct  tctctagctt  tgtgcttcatc  atggtgacca  720
gggagattc  aacatcttctt  tggaaagttg  cgacagtcaag  tatttgaag  gcgcaacgggt  790
cogtttggaa  gtaaagagc  gtgtgtaggt  gaagggcctg  gagaatttg  tcgcagcaatg  840
atctgtagc  aaaaaatctaa  agagaacact  tgtgtcagat  cccgaagag  agacaactga  900
agctaatgcg  gcgtcatctc  tctctcttcg  aagctatatc  ttcgcgaatt  tggctagtgg  960
agaaagcttg  tgcgtgtgaga  aagttcacaag  goaatttcaaa  tgcatactgac  acaagaagct  1020
gatacaatat  ccaacagcgt  ggcaatcagt  agggacatg  atgttggatt  tcctgttgatt  1080
gcagaacet  tttttttctt  aatactacaa  cggggcaaca  tgggtctcgg  ccgcaatcgt  1140
goatgtagcgc  aacagctgttg  tgtttaacaa  ttcagttgct  caagctgttt  tttaagcctt  1200
atgagtcattc  aaaaaatattc  tctctttatc  atcataaaag  acagacagc  gatgccctct  1260
ccaccccttt  gcagagcagc  ccaagtttaa  aatggagttg  aacotctata  aagctgcaact  1320
cagccctctcc  gcagaagcatg  gagagtatgc  tctcttcgcc  cgacatctcga  acctttctgg  1380
agaatatatat  cttgcagcag  gtctttttcc  ttcacatcag  goaatttcacct  tcggagtgcg  1440
cacacagcgc  gcggtacgccc  tgcagaggat  aatgtggga  gaaactcatc  aaacactcag  1500
aggggtgcgc  actcgagtgct  agaagaaactg  ccaacaatct  gcagctgctc  gcacaatctga  1560
ccaccccttg  cttgtgatgc  acggaaagaga  aactttatag  acctctctac  acaaaaaagas  1620
cgaatcgcg  ttcacaagaa  ccaatctgtat  gttactctta  aaaaaagaag  gcgtgcacaa  1680
getgacagaa  gtatcctcag  ctcgtgtcact  gcacaaccac  aagggcactt  aagagatggta  1740
tggactacct  cctttttccg  gacccatcaca  tggatgacg  aatacctgcct  ttcacaagta  1800
tgatgacact  gcacagagga  atagacact  tcccagatgt  ggtctgtgact  cogagttgag  1860
aagctgagcc  gataagcag  gttaacctgcg  aaagggcata  gttggccagc  aagaccaactc  1920
cctattagatt  ccagtgacact  cagggcagagg  caataagccaa  tgcccttaa  gatgacccaa  1980
ggggtggcaac  cagaagaca  gtcmaagagg  ccaagcatata  gggggcagac  aagacaactc  2040
cagccgcaatt  ccatactggcc  cagggcactc  cagaacaact  ccaacaagc  gttgcggcaatc  2100
ccacgacact  gacagagaa  atgaacccctc  ggctctcaacc  acggccctga  tgcggacacc  2160
aatcaagcgg  gaggagccag  cactgtgacg  tgcggacgac  gagaagttgc  gcctctgcc  2220
cttgagctca  gatgaggaag  agcagagcctg  gcagactccgtg  aaccactgtg  2280
cgcctccagc  gctccctctct  acagagatcct  ctctcttaag  aaagactctc  cgcaagacgc  2340
gcaacaagct  cagggagcccc  caggaaccctg  gcacagtgaac  aacoccaagc  2400
-continued

agaactcttc ttgaggaga tgtatgcga caattcataag tcacgagggc catttgatgc 2460
tgttttgtat tcatcataga tcgaaggatga gctctagtt ttccagtcoca gtgtgcca 2520
agagtccacg tattccacgt cccctgagg gcggatccca ccacggtcga ctggaaaaaga 2580
ggccaaggtg gaaagagaat gattgtgtaa atgggtggtg gaaacatatc atttgcgggt 2640
gataacttc aagaataaaat ctcgtgcaaat cctggacacc catcagtgaa tggacagga 2700
acacaggggt gatattcgcc agaataagcgc acaatcagct aagattaagc aatcaagggg cagaacacag 2760
aagcattgtt gattctgatat gtttatataac gccaatgccg cggagggaa 2820
aattttaccg tagctttatca ttactagccgg gttttctcaacc ttcatattgg gtcctaatgc 2880
aaaagctgc ttaagcgacttt tagaattgat caatttctgaa caatagactg aatggggtg 2940
attctaaatt aatcattcct tatttttcca aacctcctcta ctagcatgcc caacattttc 3000
agcgattgct tttaatgata aaaaaaaactg gtagaagat taaaacccttc atccctctaa 3060
cgctcaatgc aatctctgagc actgcaagct taattccttc aatttgaaaag aagaacctgt 3120
cgactcaggc caccagtcag ccaacagggc ggggccctac gctggccacg acctcaacag 3180
acacagtgcc agggtcctgg gtttcccttg ctacttctgg gcagctgtac aacggacagaa 3240
tcctgttgaag cagcattcct cttgatattg agaacaatcc aggattagtc taacacccc 3300
aaagccacaag aacggcagc aaccoctgcag cgccgcaacc tccaaacccaac agggcacca 3360
tgtcagttat gcagctttg cggctgtgc acacactca tatttctggca acctcctgaa 3420
aadacacaaac tccgctcata gaactcataag aacagacat tgaagctctg aagaatgtgc 3480
taacagcgc ttatgtatgc gcaaaaccac ttcctctttt gaccaggggt tgtggctgaa 3540
tgctgtaaa atatgtattt ctggtgtgta caccaggtcg ggtaaacagca aacgtgctgg 3600
caacctaggg tcattgtgccc gaaactctgc aaccaccacc tggacattca cttatgagg 3660
aaggtctgat cttggttgag attagaataca gagttagacg cgtctcctaa agtgggtagg 3720
agggccctaa ccacactcct aagtttctgc tcaactaaggg aaaaaatccc gggaaacactg 3780
acactctgggc aaggtacagtg aagaaacttt tcgtgatact gttggtggaac gataagtcac tcatggtaca 3840
ctctcaccgc ttttttttttt tattttttttgatatg cgggttacac 3900
tcatattgc tcgaagtccg gcagaacggc ctctactcag cttggtcctaa 3960
tccttactac aaaaaaggtt ccacacccca aagatgtgge tcaactctgc tccacacac 4020
gttcaacagtc tggactctgc cggagttgcc agaagacttt gcgtctcagtc ccaactctgc 4080
ccagagttca tggacagttt atagttcctt ttacggcctta agatggtaaa caactctgggac 4140
tcataatact gacgcaacct cctctttctc gggagaggg aataataagc gaggtctcaca 4200
cgctctatct atcggacttc taattccccg tcgagcaagc ccttggatt 4260
taataacaaaaa gctgtatttt tccatatttc tgggtacttc tctttaaat 4320
aagattgctt ggaagagtgctg cttgactaat attataacta attgttata aagtttccaa 4380
accaaaaagtt aagttgactt ggaagaaactct cactctggctg agagagtttt ttttcttaaa 4440
cctctctatct ctaaagctgg ctaatctttgtt taaaaattt gggcggtggtt atattgtca 4500
cgccctctgc tgaattatag ggatgccat acctctgcaac gttcagatc aaacattatc 4560
gggtgcaaac cagcactcag gcgctctgta cccgggcgtt agtcaagttt gcacactcat 4620
cgatcaattc gcggatgaca ccatacggca agcagaccaac acacagggga 4680
gtgtgtcaac aagcattcact cttaageta tcgtgaatgt caatacggc ccocaaagtgc 4740
taataagca aaacccaa ttggttccct tcggtgctgc tctaagaaag acctacagct 4800
ttgacataac taagccccgc atcattgctt ccatacaca tatacaccat ttggcaggg 4860
cacaacatcc atcttgacgtc gtaaatcgcg tcggctctgg attcctgact gatccctca 4920
ggtctcgag aattggaac acggttttcc tccaaggcgt cggcttccgg caagtcacac 4980
tacccagcta ttctactcct gatttgacag cactaaacct gatcaccocaa ccacagcttg 5040
cgcaacaatgc gacccgtagc acctaaacag gatcagaaag aaggtcagct ccaggaattt 5100
cattcatacc aaacattcgc cccatcttct ttcacaccca aagttggaga aaggggcaaa 5160
gtgcggattct acacattccttcg gaagaaatcc agaacaataa gacattcacc caggagttta 5220
agatggctcc aattgatcca aacccaaata tcattgagaa cstaagtcga gaaactcttg 5280
tccacacgtg gcacagcagc aagttgccct ctaaaaaatct aacacacacc atcctgtctc 5340
cttggccaaa gtcattcgag tgtgaacggg tctgtcaggg agaaccatac cattgctaca 5400
cacgagattg tgcagacggt cattcctcttg cgatgtcatt cggagtaaat 5460
tgcaataac gactcagact cagttttata gaatttcttc aggattgatg ataaactcta 5520
tttagtacg cttgacattgc agaagacact ttaaatgtat ccaactacta aaggtccctt 5580
acaacatttt ctctcttcttc tccaaatagt agaacttaaa aaaaaacata tcataatact 5640
cttctctata ccacattaagat ctaaggattc cccccatcta caccatca ctaaactacta 5700
taacaaacat ggtgattcctt atgtataacct ttttagcttc acataactctt cggctctact 5760
tccacacgtg ccatcgctctt ctcatctctgg aagctgactt atctgttctc 5820
gggaaaaat cgaggccccat ttcttaggttg tccaacatcc aagttcagcc atcctctcct 5880
ataataagaa aaacttcggtc atagagacaa acgcctcagtc gacgcctaa ttcctctctc 5940	tttagatatg ttggtgctgt acctgtgcct ttttccatgt gtaaacaaaa 6000
tacaaatcc ggaaggtcag graagggca acagcctcag gacgagttta ggaatatggc 6060
agttacttcg tgcagatgcc aagaggcaact tatcttctct ttgggctttt accttacctc 6120
aaagaacatt tccatccca ctggtgctca tccaacatag cacatccac gtattggtat 6180
tggcaaaat agtattgctg gaacaactgt ctacacaaaa taatgtgaga taagttgaaac 6240
tgatcctcga aaggggtgpa cttgcacaact gacgctcata ttagaactaa aatgaggggt 6300
tcagttcggct ttcacacca gaaagttcga attatcgacg tgctgaatgg gctgaamaact 6360
gctcaacattg tgaactccaa aaacocgagc ggaattgtgt gttcaccgca ggccgcaagc 6420
ggattgccc cttcctctcc tgcgctgtag tcgaacaaag atcgagagcg gaggtcggtg 6480
ccgagatat gcattttccttc ctttcttctc gtagatcact cttgtgtaaa 6540
cagttacca ccagggtcag actttgctct caagtcgatg tcagattctct atagcggccc 6600
aagcataagaa gacgctcattc gcggtcaagag ccagctcagc gcaagggagg 6660
acctgctag tgcacctca ttcacacca aagtaatctca ggtgcattgc ttcggtaccs 6720
atgagaagca tgtcctggtc gaggtgtcaca aattgacatg ctcacaaaat aacagatggt 6780
ttcacacca ttttgtgctc tcagctgatg agaacaata tacaagtgaggg aagaggca 6840
atacaccgag gaaactcttc tgcgaagctca aacccgacct atagcctaaa ccctgagag 6900
ggcctttctgg ggaactctaa aaaaocccac tgaacaaatt tcagctgaag tagttcttct 6960
-continued

cacagtgtta tcacacggag cccaaacaat cagtggcag agtcggcgc gaaactcttc 7020
cyacccacgg acccacaaca caactgaaga cccaaacaac agggctcag aaaaatcttc 7080
tgcaatggtt caagtgcga gtaaagagg gcgaacctga gctggtcact taaacacct 7140
tgaccaacac tccgctggag cccaatccc tccaaacca aacgacctgc acaacagcac 7200
cctataaca cccgtgtata atcttgacat ctctgtggca acaacagtgg aaccaaacac 7260
cgcagaaca gacaacgcac gcacagcttc tggacacagc cggtagccag 7320
dcccccaca gcacacagcag caagacacat gacctctgg aacccgcaac 7380
cacaccaagt ccccaaccac acgccgcac cgctgcaac acacacact aaccccaaga 7440
tacgcaagaa gagatgctca gcagcagggg gataggtta attaaacact ctatgtgag 7500
agtagcagga ctgtacacag ggccagag aactcagaga gagaacattt ccactgctca 7560
aaccnaatgc aacaccatatt tcaacactct ggttaaagct cgaagagctg ctaaagcttg 7620
actgtgctcg attaccaatt tgcgagcag aagcagagga attatacatg aggggtacta 7680
gcacaccaag gtagtttta gtgggggtgt ggaacagcgt gcacagcaga cgacacagac 7740
ttttcaacctgt ttcttcagag cccaaactga gtaacgcacc ttttaaatcc ttaaagctta 7800
ggcaattcat ttttggtgct agccagggg gcacagctgc cacattctgg gaccaagactg 7860
cgtatgcga cccacaggtt gcgaacaaag ctaacagac aaaaatctg agatactacca 7920
tgattaattct gtaaaaaacc tttggcagcc ggggagcc gcacacagtt gcacagagtt 7980
gagaacattg atacccagag gtaatgggct taagggcgtt ataatcgagc ttaacgcctt 8040
atcttctata tgcataagct ttctttaattttt tttcttcgaa tttccatcag gcacagcctc 8100
goctcaacgc aatgaaacca ggttaaatc atatagttta cttggaattca agatgacttg 8160
acaacagta ataataacac ctgtgagctt aacacagcgc aatgtgattc taaacccctt 8220
aaacaccagct ttaaataata acaaggtggg acataactct gtttacatcg tggagata 8280
taactgtgtc gaaagttaag aaaaaagttaa tttcctgatt atctttata tttattctttt 8340
attaccaatt atctgcaggt ttctttagtt acaaggggg aagcttcttt tattaagttg 8400
ataactacg atctgcagac gcggtagatt tgctgcaact ctaaacaaca taaaagcttg 8460
gtcaaaagtt cattaaaggt ttaaagcag tggagagaca actttttatg ggaagtttta 8520	tagagaag gcaccccaac agtcgcgaga cagcatcata gggagtggga gcagcagcat 8580
gttcagaca gacacatcgc cagagagat ttcaggtgg aagcagcgta ctaagggcgc 8640
goctcaacgc aatcctgcac tcacgagtt cttcagtgtt ctaaagaaga ggtatttttg tgcagtagt 8700
ccttcacgca ctaaagacat atgcggagcc tggaaaaag gattttttgt tgcagtagt 8760	ttttcagga aagcataca gaaggggagct ttaaatgata gggaattact cctttattc 8820
goctttgagc cttctggtgat cagtaacaac caatattata taacctggcc caagggcctg 8880
cggtagaaca atcacaaggg cttgtatatc cagcagacag aaggttaaaaa aattacttgt 8940	tgacacactga tcagagcgc acagacactg aagacacagc atgagggat taaatcttagt 9000
ttcaaaaattta gagacagtgg gagatctatc aaccggagac aagacatcgc 9060
cagatcgtct ttttatgtga gaccaacacta agggcgagg gccttggggc cagatcggca 9120
gcagcggcctg tctggtaata ctaacaggtt ccaactgtaa aagagcgagc cttaattagctt 9180
gccttattgc aacaatgagc cggacacatc ctaatcattgt ttacctagtc atctttgaat 9240
-continued

attgcgccg agttacgctg tgaagatgtct cttcattgg tttcagggctt aagaacacctg 9300
gttccctgg ccagataa ag gaaacccctc gccaacccctc gccacagctg cttgctccgtag 9360
gagggtaccc cttaaacc ctgacaaa ccacctataa cctctttactt gtaoacataa 9420
cctggtatcc ctcatacct atattaaact aagacagatct aaattttaac tttttacttc 9480
catatactc tctgcttca aaaaataag atgtgcctag tggcctcaat atatgaagag 9540
gtactgatac accccaacag tgaatcaggt aatacaat ctcgtatccg ctcgaatata 9600
acctgcaaga ctaactgctt gcaaaaaagt aattctgtac acaaatctgct ttttactcata 9660
caggaggtcg caacgcttcca tctccctcaaa aataattagtt ttcagactccactaacc 9720	
tttatataa taagaaaaa attccggaacc aataactctct tattgcttcaaa cttgctggag 9780
aggtgctttct ctagctataa ctgattatctt ctaataactt cccagcctagag ctggattaag 9840
tttgccattc aggtggattag tgcgacctgct aatgtaaacc taattgagaa gattatgcagc 9900
gaggtctgtat gagataaaa cttattattc agattagcgc ccacaggtcc tctcctctctctctct 9960
ccccctgcca aagactattc ctaggccttg ccaattagctt ccctgctctct ctagctgattata 10020
tcagcgccct ctagctagcc ccacaaaggt tgcctcaagcc taaattgcct gtacactcaaca 10080
ccatactatg ctaacggtcct aatatctctg tattgactc aacgtctttaa aattattgtcc 10140
atatattggt ctagacaacc caaggtccct ttccatttct ggaaagagct cattcactaaca 10200
cgaaactcct tctgcgtgtcc ctcacacaata aaaaaagga aatcagttcc aaactttcag 10260
cgcaggggttt caaggttggcc ctagagatgc ctcagacaaca aaattcagtag aactcagcaaca 10320
cacgcagag cctgcacagga aagggcctct ctagctcagc tcagcgaccc ctaactattctctct 10380
cccccaaggg cacccgacac aggggtttgc ctagactcag ctagcagaca ctaactattctctct 10440
cacactattc aaggggtctga gttttactgg cgtcctgttg agtttttttag gtacccacaca 10500
gagatggcct ctagcagtag tgcgaacatt actgaattgg cccctgcatag ctaacacaca 10560
aggaatcttg tctccctgcat ttttcaaat cgaatttcca caaatgataccc accgtcttgagg 10620
goatcagc tcctccgcttg acagaggtata cagacggctc ctagacgctgc tgcctgtatt 10680
gagccgcttg agggcccttt cgtgcgttcc cttgcagcctactg tccctctctag cagcactcaca 10740
catctcaca ccagcgaccc aacgtctcag ccagactaaat gtgcgcctt ccctctcctg 10800
atttccatc ctagtccaaa gtccattacat aatagtggatt cttcctactgt gattacactg 10860
aacaggttgt tcagcagactgt cagcaattgaa atccaaaactt ttctgctctt aaattgacgctg 10920
acactattcg gttttacttg ggaaccttaca aacgcgcatc ccagcagtctg gacccagcatc 10980
aagctctg gccgagaaat ttttcccctct tagagatggc ccagctgaaac atttacacaa 11040
sggtccggc ccaagctacca aagttttgat tttttactct ctagacattc ttcatctttttg 11100
catatattag atactctctag ctagcagatac tcctcttcat ttctgcctc gttatccttg 11160
cacatttcag tttctctctag cttttatat cagcttacta aatgttccg ttcctgtatctg 11220
aatattttttt ccaacagactt gattctctctt ttccagctag ccccttccttc ctaacagtctc 11280
aaaaacontg ccagcactag ctctcttcct gttaaacgtgc ttcttcagttt aacccctttc 11340
attttactag atacagaggct ttcaccacag ccaaaacctg tttacaagtt ttagttgtgaa 11400
cctctctgctg caggtcagctgt gtttagttg ccaagctttct atatcatctg aacccacttc 11460
togatactt cattttatat gttagattcag aaaaaagcct gggaaagttt aagaaaaact 11520
-continued

ggtgacactc agtgacattc tggattatac gtcttcccct tagagactga gcagacgcag 13860
cagacaaga gcggcagaia caatgcagcg aggggtgccg cccgacctac aagatctaca 13920
agtgctgttg gaaatctttt aaaaaccttg gaaacatttg taacattcagg ttttatctat 13980
tttgaaaaa aacactattt gattggggtc caattgcttc ctggcccttaa aacggttaca 14040
agaatggtcgc caattggctga tcgatttttt tattgatctta aagggacctt ggcatagtata 14100
rgcactgcttt tgtagcagtc caatttgtag acacagacata tttttctttg cagagatacc 14160
gcgatttccc atatcttttt ttcgggtgaga atctttcgat attactacatt cgggctcaat 14220
aagagtttttg actctgtgaca gttacacact ggcaacacte tggattcagg aacaaatatc 14280
ttgctgtaaat gcgttctgga gggattacct ttctgaaact tcagaaattg 14340
ttctacgaga atcttggatt gacactgatt gtcagtttat tcatagttaaa aacctaatctc 14400
cgaaatatcg aactggtgat gttattcctta otctttatgt ccaagaaacc gggagagcct 14460
actcgaattg acctttggtct aattcctcag gcagtaatga ttcctggtgct gcagaacctta 14520
actctatctc tggccgctatgt gtcagctcagct acacccctcc aacatgccaa aacaaactttc 14580
attacatcct tatttcctag gtcagctcagct ttcgagagct aatggttttg taatggtcct 14640
ttatcatcaaa ctctctggatat ggctgtttttt gcggctgatata ttttccctagc cgcagccgagc 14700
gggaaagcgt tggatatcttc aagagaactg gcggagaccc caagactctg agcctctag 14760
atctcaca acatgctgac gacccctggg tttgacacgc tgacagaaat aacattgcaac 14820
aggggtgagcc tgttggtagag tttttcttgat ctattctgtttag gattgtgatat 14880
accacaaaaa ctttacgtct gattttctgaa ccagtttttg ggaggtatctt atgggctcat 14940
attttagg aggregcctct attttggagcc aacactacat gtaattgtg gcacagtaaa 15000
gtttttggc tgaaaccccta caagaaatgt tgcagcttggg caatccagtaa gcagaccaggtt 15060
gggaac’accattggtgtggtgcaag aacaatattg ttaagctgctg gcagaagcctgtaa 15120
ttcgacatccag cttcgctatt gcggagttcag accctatcaaa ttgctttccagc gcagagaaat 15180
aagatagcgg aaccctgtatt taaacaaaaa tgtctttctgg cagcctttaa aagggccctt 15240
gaattggcct ccctgtttttc atggcttaact ccagccagttc caagcactgtg cccttgctaat 15300
aaccattttt tgggaagcagagaatattga aagttggctag aacactctca aatgacccct 15360
tccattact caggaatatttgttcagctt cacaactcagctatcattcttctgc 15420
atggccagac tgtgagctaa ttcagcaacc gcctttgtggt ttcatctaaaa caacatggtg 15480
gatgttggcc aggggttggcag ttcgacgcc gcacagcataa ttttctttctgaa gaagttttat 15540
aatatgctgg tgtgcgctttt gctatattaa tttggagaaa atctgatcagc agatctccaa 15600
atatattctg tctcagcttc taataagctgc tttgctcccgg gggagatctt gtcgagctt 15660
atggccgctgtttt tacttcattg gattttgat ttaaagaagct aagagacagg caagatgtg 15720
attgaggcct atcctcctaca agggagcttaa aagctctcata ttcattctaat gtaatcattt 15780
cttcagagaa aucgcttga cattataaa gcctagcttt ttcgagctcgg ccaatcatttt 15840
gagatgggagc tagctgaacag cattctcctta tcagttattt cgcatacgca caattctctc 15900
acagagccaa ttcagctgctgg agaaaaacagata ctcattctct aacctttctt acatattcacc 15960
aagtaggtcgt ttcgctcgca tttttggttgg cttctttgct tattatcttg ccagttcact 16020
cgctgacta agaaaactacc actctgacaat ttttattactt acttaactac ttcattttcat 16080
-continued

aacctacca c atcgctcatt gcgaatacct aacgccaacat tcaaacactgc aagcggttag 16140
tcaaggttta tggatttga ttctcatttt ctatttaca taggctggcgt tgcaggtggc 16200
agaggactct cagatggcgc caggttatttt ttgagaacgt ccatctttcat ttttottaca 16260
ttttgaaag aatggtatgt ttaagcggga aacaattgcc cttatggat aagtataccg 16320
ttagggtgct aaaaacccaa acctgtgaat aatitmitct acatatagct agaactgtgc 16380
tgtcatgtatt ccatcaagca cagggctttt aaaaactacca taagtgatca tgcatact 16440
cacgacaact tgggttctac aagtaagagt aacgccccga aacctcctca tgcagctattg 16500
gcgtactcga ggagcagaca cagaaacagc acacggcatt acctgagcaag aagccttctca 16560
acattgatcga gccaaacaca caggtgttgt cattttgaga gaaactaaga acaaaaccac 16620
agagatocac atgtagggcgc tgaacggatct atggctctac aattagcagc gaaataaaaa 16680
agaaagcgcct ttcgcacaga aacaaacbtc acgggtctcg cgtttccagtc cttccctaag 16740
gctctgcttt tgggtacagc aatcccaaata ctaaatctcg atcgatcagc aacaataagcgt 16800
aatttccagg atatcaactg gctataccag agggagttag atcaatataa atccacccagt 16860
ttagctctca attttttcaaa attatttcga gggacagcc aatttaagctg atccaatagag 16920
tcaacacaa aagacgtgat atcagatcct ctaggctcag ctagttacagc 16980
acagtttatt gtgatgttca cgtgattagc tccctctgagc attcaacactc ttaggggttc 17040
tttttgaga aacaaaccag cagggcagct cagttggcttcgg cagggcaggag tgggtccg 17100
gacattctat tgtgatcagag ataaactttat tttactacaag cgtotctcttg 17160
gatgtcgagtc tgaactcagac aataagttta gggagcttc aatgccgggt accaagcag 17220
gttatggcag aattaccatgg cgagattctc ttaaactaag ttcagagca 17280
atcaagcaac tcaacaacccc tacattggttt atacgcaactaatg gacaaactcat 17340
gtcgggtga taacacagccag cagacgagaa aatcaagactg aatctctgac 17400
gagctgttat aataaatcctg ctttacacat atggatcctt cgtctttgaga aacaggtcgc 17460
cctttacttt cctataagtgct cagttgatgc attaatgataa tttacctcctg 17520
cccccccctc ttgttttttt tttgtaaactct caggtgtcagag acacgtaaag ttcaccacacc 17580
tagtgtaac agggaaactct cagggcattag caggtgttct cggcgatctg 17640
cttaggtcct ctaccaacgtag tgcgcctcg atcagagagac ctacatatggat 17700
ctagttgcc aatctcactag tgcgctaggct gattttctctt cagttgctag tgggt 17760
tcctacatct taggaagagc agtatatacc tctaggttca cagttgcaga taaagctttt 17820
cctcagttct ctaccaactct gcaatttctag caggttagag cagatgactag 17880
atatgattata aatacagccg acagaaatccg aataacactt caaatctttt ttcgctg 17940
aaagcacgaa ttcacaaact aagctacagt tatttaaaat ttttttctttt ttgagcagca 18000
ttttattttttt cctgagcttt ttcggaatat tccagagagc gatatttgtg 18060
agagagttagt tcttcataat tagagattgc ctagttgagag aacgtttcttc tgtcctacac 18120
atatattttttt atacagttcag gatttcttttt cgcagttcag cagggggttt 18180
cctgaattct ttcgctgatgc ttctctacgtg ttgaggctgta ttagacagtc 18240
atatsatcgaag aggctgttatg ttcacacagcac ctttcgctg 18300
acaccaatttt gattctttt gtaaatcaac ccattttatt ttttaaatc tttgttaatttt 18360
tataatacat attagggttcct tcctccctcgt gcgatactcaaa aaaaattca ca atgcagctag 18420

tgtgacatag tacattcgca atgaatttaa cgacacataa taataactctgc acottttata 18480

attaagcttt aacgaaaggt ctggggtcctat atgtgattag atataataaatgttgcataca 18540
taccctgctca gatggcagatcc ttttctggtttt gataacacaa cttctctaaataa caaatattgt 18600
ccttaaggatt aagttttttta taattatcacat taccttaaatctgctgtttta aaaaaggtgta 18660
tagcctttaattttgttaaa aataagagatatggcgtaat caaccttaacat ttttctgctca 18720
gtaagctct attttatcata gaattgataaa atttaaaagaa aagccagagc tggtaaatac 18780
gaataacctc cttaacataa tagcagactata taatataatact cttgcgtttaa tgataatataa 18840
gacattgaco acgctcaaaa gcagctctgc cagaataaaa ggtcgcattgg cc gatttctcg 18900

aaaaagttgc gcacaaaaaa attttaaaaat aatctctatttt cttctttttttt gttgcgtcaca 18969

<210> SEQ ID NO 21
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Sudan ebola BMG
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8)..(8)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (8)..(8)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 21

gccatgnttt caggttgtag

<210> SEQ ID NO 22
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Sudan ebola BMG
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: I
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: n is a, c, g, or t

<400> SEQUENCE: 22

ggtacatctg gcacacattt ca

<210> SEQ ID NO 23
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR probe for Sudan ebola BMG
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(1)
<223> OTHER INFORMATION: Fluorescein (FAM)
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (26)...(26)
<223> OTHER INFORMATION: Black hole quencher dye (BHQ1)
<400> SEQUENCE: 23

aeggtgoaca ttotcctttt otogga

<210> SEQ ID NO: 24
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment A

<400> SEQUENCE: 24

gtgagacaaa gacaccctccc ccg

<210> SEQ ID NO: 25
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo fragment A

<400> SEQUENCE: 25

catcaattgc tcagagatcc acc

<210> SEQ ID NO: 26
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment B

<400> SEQUENCE: 26

ccacaacacc tgcagtagaag t

<210> SEQ ID NO: 27
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo fragment B

<400> SEQUENCE: 27

aggtgcggtt asctcccatac

<210> SEQ ID NO: 28
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment C

<400> SEQUENCE: 28

gatgggttac ttaacttcccag g

<210> SEQ ID NO: 29
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo
fragment C

<400> SEQUENCE: 29
gtttagtct atcaatgcc

<210> SEQ ID NO 30
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment D

<400> SEQUENCE: 30
ccaccagcac caaagcac

<210> SEQ ID NO 31
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo fragment D

<400> SEQUENCE: 31
catcgccaa tgaactatt gg

<210> SEQ ID NO 32
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment E

<400> SEQUENCE: 32
gcctttgtag aggacacac

<210> SEQ ID NO 33
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo fragment E

<400> SEQUENCE: 33
ccacattaat tgttctaca tgcaag

<210> SEQ ID NO 34
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Ebola Bundibugyo fragment F

<400> SEQUENCE: 34
cctagtgttatt tagaagggcta

<210> SEQ ID NO 35
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
-continued

<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Ebola Bundibugyo fragment F

<400> SEQUENCE: 35

gtagatga ttgacagca ta

<210> SEQ ID NO 36
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer for Ebola Uganda 692(-)

<400> SEQUENCE: 36

acaaaaagct attgtcacta t

<210> SEQ ID NO 37
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR primer for Ebola Uganda 18269(+)

<400> SEQUENCE: 37

ctcagaagca aatattaatgg

<210> SEQ ID NO 38
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote diIvoire ebola virus fragment A

<400> SEQUENCE: 38

gtgtgagaat aactatggag aag

<210> SEQ ID NO 39
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote diIvoire ebola virus fragment A

<400> SEQUENCE: 39

gtctgtgcaaa tgtgtgagaa gg

<210> SEQ ID NO 40
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote diIvoire ebola virus fragment B

<400> SEQUENCE: 40

catgaaac acacacaaca ac

<210> SEQ ID NO 41
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoire ebola virus fragment B

<400> SEQUENCE: 41

gttgcttaa tttcatcaatgt

<210> SEQ ID NO 42
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote divoire ebola virus fragment C

<400> SEQUENCE: 42

ggctataagt aatcttccttc

<210> SEQ ID NO 43
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for ebola cote divoire virus fragment C

<400> SEQUENCE: 43

caggtgtatt gttgtgctta gc

<210> SEQ ID NO 44
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoire ebola virus fragment C

<400> SEQUENCE: 44

gctggcattg gaacacaggg

<210> SEQ ID NO 45
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoire ebola virus fragment D

<400> SEQUENCE: 45

cgtgtagcta cagttcttta g

<210> SEQ ID NO 46
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote divoire ebola virus fragment E

<400> SEQUENCE: 46

gacaaagga ttagttagcg tatag

<210> SEQ ID NO 47
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoiré ebola virus fragment E

<400> SEQUENCE: 47

tgaatgagaa ggtgctaatgg

<210> SEQ ID NO 48
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote divoiré ebola virus fragment F

<400> SEQUENCE: 48

cagcacttag ttgacataatt gg

<210> SEQ ID NO 49
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoiré ebola virus fragment F

<400> SEQUENCE: 49

cagcactacta ttgatctgg aag

<210> SEQ ID NO 50
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote divoiré ebola virus fragment G

<400> SEQUENCE: 50

cgacacaca aaaaagawra a

<210> SEQ ID NO 51
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote divoiré ebola virus fragment G

<400> SEQUENCE: 51

cgtctcgag cttagcagtt c

<210> SEQ ID NO 52
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote divoiré ebola virus fragment H

<400> SEQUENCE: 52

gcaactataag ctcgatgaag tc
<210> SEQ ID NO 53
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote d'Ivoire ebola virus fragment H
<400> SEQUENCE: 53

tgscacaca aaargaraa

<210> SEQ ID NO 54
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for Cote d'Ivoire ebola virus gap between fragments C and D
<400> SEQUENCE: 54

cgtsgagagc ccgsagaaag g

<210> SEQ ID NO 55
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR reverse primer for Cote d'Ivoire ebola virus gap between fragments C and D
<400> SEQUENCE: 55

gtsgacgt tgtataoct cc

<210> SEQ ID NO 56
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for ebola uganda virus EboU965(+)
<400> SEQUENCE: 56

gsgaaaggg cgtctcggag aa

<210> SEQ ID NO 57
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR forward primer for ebola uganda virus EboU1039(-)
<400> SEQUENCE: 57

tgggattcg aatcagacct tgtt

<210> SEQ ID NO 58
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: PCR probe for ebola uganda virus EboU989
<400> SEQUENCE: 58

ttcaacaga aatccagagt cacgca
<210> SEQ ID NO 59
<211> LENGTH: 302
<212> TYPE: PRT
<213> ORGANISM: Bundibugyo ebolavirus
<220> FEATURE:
<221> NAME/KEY: misc_feature
<223> OTHER INFORMATION: SSGP viral protein

<400> SEQUENCE: 59

Met Val Thr Ser Gly Ile Leu Gln Leu Pro Arg Glu Arg Phe Arg Lys
 1   5   10  15

Thr Ser Phe Phe Val Trp Val Ile Ile Leu Phe His Lys Val Phe Pro
 20  25   30

Ile Pro Leu Gly Val His Asn Asn Thr Leu Gln Val Ser Asp Ile
 35  40   45

Asp Lys Leu Val Cys Arg Asp Lys Leu Ser Ser Thr Ser Gly Leu Lys
 50  55   60

Ser Val Gly Leu Asn Leu Gly Asp Gly Val Ala Thr Asp Val Pro
 65  70   75  80

Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val
 95  90  95

Val Asn Tyr Glu Ala Gly Glu Trp Ala Glu Asn Tyr Asn Leu Asp
100 105 110

Ile Lys Lys Ala Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly
115 120 125

Val Arg Gly Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
130 135 140

Gly Pro Cys Pro Glu Gly Tyr Ala Phe His Lys Glu Gly Ala Phe Phe
145 150 155 160

Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Ser Thr Thr Phe
165 170 175

Ser Glu Gly Val Val Ala Phe Leu Ile Leu Pro Glu Thr Lys Lys Asp
180 185 190

Phe Phe Glu Ser Pro Pro Leu His Glu Pro Ala Asn Met Thr Thr Asp
195 200 205

Pro Ser Ser Tyr Tyr His Thr Val Thr Leu Asn Tyr Val Ala Asp Asn
210 215 220

Phe Gly Thr Asn Met Thr Asn Phe Leu Phe Glu Val Asp His Leu Thr
225 230 235 240

Tyr Val Glu Val Leu Glu Pro Arg Phe Thr Pro Glu Phe Leu Val Glu Leu
245 250 255

Asn Glu Thr Ile Tyr Thr Asn Gly Arg Arg Ser Asn Thr Thr Gly Thr
260 265 270

Leu Ile Trp Lys Val Asn Pro Thr Val Asp Thr Gly Val Gly Glu Trp
275 280 285

Ala Phe Trp Glu Asn Lys Lys Leu His Lys Asn Pro Phe Lys
290 295 300
1. An isolated hEbola virus comprising a nucleic acid molecule comprising a nucleotide sequence of:
   a) a nucleotide sequence set forth in SEQ ID NOS: 1 or 10;
   b) a nucleotide sequence hybridizing under stringent conditions to SEQ ID NOS: 1 or 10;
   c) a nucleotide sequence of at least 70%-99% identity to the SEQ ID NOS: 1 or 10, with the proviso that said nucleotide sequence is not SEQ ID NO: 20.
2. An isolated hEbola virus having Centers for Disease Control Deposit Accession No. 200706291.
3. The hEbola virus of claim 1 which is killed.
4. The hEbola virus of claim 1 which is an attenuated hEbola virus.
5. The virus of claim 4 wherein at least one property of the attenuated hEbola virus is reduced from among infectivity, replication ability, protein synthesis ability, assembling ability or cytopathic effect.
6. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NOS: 1 or 10 or a complement thereof, or a fragment thereof wherein said fragment comprises a nucleotide sequence of between 4 and 4900 contiguous nucleotides of the nucleotide sequence of SEQ ID NOS: 1 or 10, or a complement thereof with the proviso that said nucleotide sequence is not comprised by the nucleotide sequence set forth in SEQ ID NO: 20; or between 5500 and 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NOS: 1 or 10, or a complement thereof.
7. The isolated nucleic acid molecule of claim 6 comprising a nucleotide sequence of between 4 and 4900 contiguous nucleotides of the nucleotide sequence of SEQ ID NOS: 1 or 10, or a complement thereof; with the proviso that said nucleotide sequence is not comprised by the nucleotide sequence set forth in SEQ ID NO: 20; or between 5500 and 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NOS: 1 or 10, or a complement thereof.
8. The isolated nucleic acid molecule of claim 7 comprising a nucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 2-9, 59, or SEQ ID NO: 11-19 or a complement thereof.
9. An isolated RNA or DNA nucleic acid molecule which hybridizes under stringent conditions to a nucleic acid molecule having the nucleotide sequence of SEQ ID NOS: 1 or 10 or a complement thereof.
10. An isolated polypeptide encoded by the nucleic acid molecule of claim 7.
11. The polypeptide of claim 10 comprising the amino acid of:
   a) an amino acid sequence set forth in any of SEQ ID NOS: 2-19, or 59; or
   b) an amino acid sequence that has 70%-99% homology to the amino acid sequence of (a).
12. The polypeptide of claim 10 wherein the amino acid sequence has
   5 to 250 contiguous amino acid residues of the amino acid sequence of SEQ ID NOS: 5 or 18 (VP24);
   5 to 280 contiguous residues of the amino acid sequence of SEQ ID NOS: 6 or 17 (VP30);
   5 to 320 contiguous residues of the amino acid sequence of SEQ ID NOS: 8 or 13 (VP40);
   5 to 340 contiguous residues of the amino acid sequence of SEQ ID NOS: 7 or 12 (VP35);
   5 to 370 contiguous residues of the amino acid sequence of SEQ ID NOS: 4 or 15 (SGP);
   5 to 370 contiguous residues of the amino acid sequence of SEQ ID NOS: 59 or 16 (SSGP);
   5 to 670 contiguous residues of the amino acid sequence of SEQ ID NOS: 9 or 14 (GP);
   5 to 730 contiguous residues of the amino acid sequence of SEQ ID NOS: 3 or 11 (NP); or
   5 to 2200 contiguous residues of the amino acid sequence of SEQ ID NOS: 2 or 19 (E).
13. (canceled)
14. (canceled)
15. (canceled)
16. (canceled)
17. (canceled)
18. (canceled)
19. (canceled)
20. The hEbola virus of claims 3 or 4, or a protein extract thereof, and a pharmaceutically acceptable carrier.
21. (canceled)
22. The nucleic acid molecule of claims 6 or 9, and a pharmaceutically acceptable carrier.
23. (canceled)
24. (canceled)
25. (canceled)
26. (canceled)
27. (canceled)
28. (canceled)
29. (canceled)
30. (canceled)