SIANN: Strain Identification by Alignment to Near Neighbors

Samuel S. Minot, Stephen D. Turner, Krista L. Ternus, and Dana R. Kadavy

January 9, 2014

Abstract

Next-generation sequencing is increasingly being used to study samples composed of mixtures of organisms, such as in clinical applications where the presence of a pathogen at very low abundance may be highly important. We present an analytical method (SIANN: Strain Identification by Alignment to Near Neighbors) specifically designed to rapidly detect a set of target organisms in mixed samples that achieves a high degree of species- and strain-specificity by aligning short sequence reads to the genomes of near neighbor organisms, as well as that of the target. Empirical benchmarking alongside the current state-of-the-art methods shows an extremely high Positive Predictive Value, even at very low abundances of the target organism in a mixed sample. SIANN is available as an Illumina BaseSpace app, as well as through Signature Science, LLC. SIANN results are presented in a streamlined report designed to be comprehensible to the non-specialist user, providing a powerful tool for rapid species detection in a mixed sample. By focusing on a set of (customizable) target organisms and their near neighbors, SIANN can operate quickly and with low computational requirements while delivering highly accurate results.

Introduction

There are many different methods that characterize the mixture of organisms present within a metagenomic dataset. Such datasets are generated when a complex environmental sample is processed by a “next-generation” high-throughput genome sequencing protocol, and they consist of large numbers of short nucleotide sequences. Each sequence represents a small fragment of a randomly selected genome from the very large collection of genomes present in the source sample. Those sequences indicate the presence of one organism or another according to their similarity to a set of known reference genomes. While a given sequence may be unique to one species, it also may be found in diverse organisms across the tree of life. Therefore, one analytical challenge (among many) is to take that collection of sequences (likely numbering in the millions) and accurately determine what species are present in the sample. Here we describe a novel method (SIANN: Strain Identification by Alignment to Near Neighbors) that is specifically designed to rapidly detect a set of targeted organisms from a metagenomic dataset by aligning reads to genomic regions that are unique at the strain or species level.
SIANN: Strain Identification by Alignment to Near Neighbors

The analytical question motivating a particular piece of metagenomic bioinformatic analysis may vary widely by user and sample type (Segata, et al., 2013). For example, the function of the human gut microbiome may depend on the relative abundance of hundreds of species of bacteria and the types of metabolic genes they contain (Wu, et al., 2011; Schloissnig, et al., 2013). In contrast, the clinical treatment of a patient may depend on whether or not a particular virus, or a consortium of co-infecting pathogens, is/are detected in their blood. It is this second class of presence/absence questions that SIANN is designed to address. SIANN is appropriate for situations in which a user wants to know whether a particular organism or set of organisms is present in a sample, but isn’t interested in the functions encoded in their genomes, the relative abundance of each organism, or any other more in-depth analysis.

Methods

Approach
Metagenomic classification methods are based on a wide variety of theoretical underpinnings. The basic varieties include alignment of reads to various nucleotide databases or exact matching to nucleotide or protein signature sequences (or kmers). A representative set of recent methods are described in Table 1 (also see Bazinet & Cummings 2012).

<table>
<thead>
<tr>
<th>Name</th>
<th>Method</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEGAN</td>
<td>Alignment to large nucleotide database</td>
<td>Huson, et al., 2011</td>
</tr>
<tr>
<td>PhymmBL</td>
<td>Alignment to large nucleotide database with interpolated Markov models</td>
<td>Brady &amp; Salzberg, 2011</td>
</tr>
<tr>
<td>MetaPhyler</td>
<td>Alignment to clade-specific marker genes</td>
<td>Liu, et al., 2011</td>
</tr>
<tr>
<td>MetaPhlAn</td>
<td>Alignment to clade-specific marker genes</td>
<td>Segata, et al., 2012</td>
</tr>
<tr>
<td>LMAT</td>
<td>Nucleotide kmer matching</td>
<td>Ames, et al., 2013</td>
</tr>
<tr>
<td>Kraken</td>
<td>Nucleotide kmer matching</td>
<td>Wood &amp; Salzberg, in submission</td>
</tr>
<tr>
<td>Sequedex</td>
<td>Protein kmer matching</td>
<td>Berendzen, et al., 2012</td>
</tr>
<tr>
<td>mOTU</td>
<td>Alignment to universal marker genes</td>
<td>Sunagawa, et al., 2013</td>
</tr>
<tr>
<td>Phylosift</td>
<td>Insertion into reference nucleotide and protein alignments</td>
<td>Darling, et al., in preparation</td>
</tr>
</tbody>
</table>

Table 1. Summary of methods for metagenomic classification.

Overall, these methods are designed to either classify individual reads to, and/or predict the total abundance of, clades (e.g. genus or species) across the entire tree of life. They generally require reference databases that are very large and/or require a large amount of processing to generate. The gap SIANN is designed to fill is when the entire tree of life is irrelevant, and only predefined subsets of organisms need to be detected. For an underlying method we chose read alignment to diagnostic genomic regions because the algorithms for read alignment are highly parallelizable and have been optimized heavily by the community at large (the current implementation of
SIANN: Strain Identification by Alignment to Near Neighbors

SIANN uses bowtie2 [Langmead & Salzberg, 2012] for the alignment function, but can be adapted to any alignment algorithm. This approach is distinct from using clade-specific marker genes (Segata, et al., 2012) because unique regions that are larger, smaller, or outside of genes can also be used. Furthermore, this approach supports the rapid construction of custom databases using reference genome sets that require only minimal user-supplied structure.

To understand the principle at work, consider a set of reads that have been aligned to the genomes of several strains belonging to two species. Some regions of those genomes are species-specific, some are strain-specific, and some are shared (Figure 1a). When a set of reads is aligned to those genomes such that each read is placed in as many locations as it has a match (at a reasonably stringent threshold), visual inspection of the distribution of reads yields an intuitive understanding of the true source organism as Species I/Strain B (Figure 1b). If Strain B were not present in the reference database, it would still be clear that the organism was an unknown strain of Species I.

The unique identification of a species or strain is quantified by the proportion of the genome that is determined to be species- or strain-specific (defined as reads that are aligned to regions that are species- or strain-specific). Each species and strain is then assigned a numerical measure of the proportion that is covered by these diagnostic reads, and that proportional measure is compared to the ideal case, where sequences from a single organism (generated in silico) are aligned against the database in an identical manner. After that normalization factor is applied, the resulting score indicates whether the source sample contained any of the organisms in the reference database.
SIANN: Strain Identification by Alignment to Near Neighbors

The analysis is conducted independently on both the species and the strain level, so that if the true strain is not present in the database, the species of origin will still be identified. While many methods consider the complete taxonomic tree and assign reads to the least common ancestor, SIANN considers only two taxonomic levels: species and strain, throwing out anything that is not unique at one of those levels and thus obviating many of the confounding factors introduced by manually curated taxonomies.

The example shown in Figure 1b indicates that species-specific reads are identified as reads that align to one species (Species I, in that case) but not the other. If Species II were not present in the example shown in Figure 1b, a much larger number of reads would be assigned as “species-specific,” when in fact those regions are shared with other species. Therefore, the ability of this method to identify strain- and species-specific sequences is a direct function of the inclusion of near neighbors in the reference database. This characteristic is shared among many classification algorithms, but it is of particular note for this method when users have an opportunity to construct their own database. In order to detect a target species with a high degree of specificity (reducing false positives), it is necessary to include other related species in the reference database. Only by parallel alignment to those near neighbors can the redundant sequences be separated from the species-specific ones. For example, in order to detect Bacillus anthracis in a sample, it would be necessary to include other species of Bacilli in the reference database so that the presence of B. cereus or B. thuringiensis in a sample does not lead to a false call for B. anthracis.

The nomenclature of genus, species, and strain is potentially problematic because it does not correspond to a consistent degree of evolutionary distance or genomic distinctiveness. The ability to distinguish two organisms by any method using genomic sequence data is proportional to the amount of each genome that is shared or unique. One might assume that any two organisms of the same species will have a relatively predictable amount of shared genomic identity. However, some pairs of organisms from the same species may have less in common than other pairs of organisms from different species or even genera. This ambiguity impacts SIANN in two ways. If two organisms have very little genomic sequence to distinguish them, the sensitivity of SIANN to detect either one will diminish (the rate of false negatives will increase as the likelihood of sequencing unique regions decreases). Conversely, if an organism is extremely dissimilar to the near neighbors selected for the database, the specificity with which SIANN detects that organism will decline (the rate of false positives will increase as the number of related genomes available in the database decreases). For example, if a database contained only E. coli and B. anthracis, a sample containing B. cereus would be misidentified as containing B. anthracis. In the intended use case, a database targeting B. anthracis would contain B. cereus and a number of other near neighbors to prevent that kind of misidentification. It would be convenient to say that
SIANN: Strain Identification by Alignment to Near Neighbors

An ideal database can be made by calculating the ideal genetic distance between all references and then finding an ideal set of organisms to make up that database, but the behavior of any database will be governed by the particular genomes of the organisms it encounters in the wild. Because not all organisms evolve in the same manner (differences in mutation rate, horizontal gene transfer, recombination, etc), the suitability of a database and method to detect a given organism can only be determined by thorough validation and benchmarking, as well as updating the reference database as needed. Users of SIANN may construct their own custom databases to include newly identified genomes or specific subsets of genomes that best suit their research interests.

Steps to construct a custom database:

1. Select a set of target organisms
2. Gather a set of genome sequences for those target organisms as well as a matched set of near neighbors
3. Using those reference genome sequences as an input, SIANN will:
   a. Construct a reference index for alignment
   b. Simulate a set of reads from each genome
   c. Align each of those simulated read sets to all of the reference genomes
   d. Calculate the proportion of each reference genome that is strain- or species-specific
   e. [If two organisms do not have a minimal amount of unique sequences that exceeds the rate of sequencing error, SIANN asks that all but one of those organisms are removed from the database to eliminate redundancy. Note that the user can provide a single representative genome with multiple strain names so that the redundant strain names are not lost.]

The files contained within each SIANN database are a compressed genomic index and a list containing the proportion of each reference genome that was found to be strain- or species-specific during database construction.

To run SIANN:

1. Select a pre-made SIANN database and a set of sequences to be analyzed, and
2. SIANN will:
   a. Align each of the reads against the reference genomes
   b. Calculate the proportion of each reference genome that is strain- or species-specific within those reads
   c. Compare that proportion to the simulated ideal case generated during database creation
d. Calculate the probability that the given results could be generated by random chance

e. Report the normalized proportion and non-parametric statistic of likelihood for each strain and species in the reference database. The normalized proportion of the genome covered by strain- or species-specific reads is the primary statistic reported by this tool.

**Benchmarking**

The performance of SIANN (version 1.6) was tested in comparison to the following state-of-the-art metagenomic classification programs: LMAT (version 1.2), MetaPhlAn (version 1.7.7), and Kraken (version 0.9.1b). All of the programs in Table 1 were investigated for this effort, and three were chosen based on their ability to run on our high-performance computing cluster with an execution time and memory requirement that would be suitable to a clinical lab. Each program was run on a set of 600 simulated datasets generated by MetaSim (Richter, et al., 2008). Each dataset consisted of 15,000,000 reads (100bp single-ended) with Illumina-simulated error (fourth-degree polynomial) (Korbel, et al., 2009). The 600 datasets were broken into 12 sets of 50 replicates. Each of the 12 sets contained organisms at different levels of abundance as shown in Table 2. Organisms were specifically chosen in pairs so that the ability to distinguish these near neighbors could be determined. The abundances were staggered at 4-fold intervals so that a wide range could be evaluated. All known species of near neighbors for each of

<table>
<thead>
<tr>
<th>Organism</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Bacillus anthracis</em></td>
<td>0.074%</td>
<td>0.30%</td>
<td>1.2%</td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Bacillus cereus</em></td>
<td></td>
<td>0.074%</td>
<td>0.30%</td>
<td>1.2%</td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Hanta virus</em></td>
<td>1.2%</td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
<td>1.2%</td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
</tr>
<tr>
<td>Rift valley fever virus</td>
<td>0.30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19%</td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
<td>1.2%</td>
<td>4.7%</td>
</tr>
<tr>
<td><em>Clostridium botulinum</em></td>
<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><em>Clostridium difficile</em></td>
<td>1.2%</td>
<td>4.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Listeria fieschmann eii</em></td>
<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><em>Listeria monocytogenes</em></td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monkeypox virus</td>
<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><em>Vaccinia virus</em></td>
<td>19%</td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Yersinia enterocolitica</em></td>
<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><em>Yersinia pestis</em></td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
<td>1.2%</td>
<td>4.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. The abundance of each target organism in each set of simulated datasets. Each set is indicated by the number in the top row, and was generated with 50 replicates.
the 12 target organisms were included in the reference database used by SIANN for this benchmarking (“Target Pathogen Database”) and are shown in Appendix 1.

Each program outputs a distinct measure. Kraken and LMAT both count the reads assigned to each taxon, MetaPhlAn calculates the abundance, and SIANN outputs a measure of the proportion of diagnostic genomic regions present. To put these measures on an even footing, we empirically calculated the false positive rate for each method over all 600 samples, at each possible measure of output. Because each dataset is made up of known organisms, any result can be classified as true or false. Therefore, for any possible result (say, 513 reads classified by LMAT or 1.6% abundance assigned by MetaPhlAn), one can calculate the proportion of calls with at least the same amount of support that were correct (True Positives/[True Positives+False Positives]), over all of the 600 datasets. That measure is commonly given as Positive Predictive Value (PPV). For each program, the results can be translated from the raw value into a PPV that is based on this empirical measure of error. The key item of interest is the PPV value for the results that we know to be true positives, the defined spike organisms. Another way of describing this approach is to say that the results of each program have been normalized to the false positive error rate that was empirically observed. If another set of samples were generated, the PPV vs. raw value curve (Figure 2) would likely fall differently, but in this case it gives us a means of comparing a diverse set of methods against the same ground truth. If method 1 detects an organism with a higher PPV than method 2 does, it means that method 1 has fewer false positives in the range that it reports true positives, which is the definition of utility in this setting.

For each method, PPV was calculated as a function of raw output value. Briefly, this was done by compiling the output for all 600 samples, labeling each result as false or true based on the sample set that it came from, and then calculating (at each possible value of output) what the proportion of TP/[TP+FP] was for results with at least that level of raw output. Some simplification steps were taken, such as focusing on the species-level assignments (for comparison with methods that do not perform strain assignment), and only taking the top hit for each species from each dataset. Custom R and BASH scripts were used for the data compilation and analysis.

Results
The relationship of raw output value to PPV is shown for each of the four methods in Figure 2. The point at which PPV is very close to 1 (where 95% of results are true positives) is ~41,000 reads for Kraken, ~2,800 reads for LMAT, ~38% abundance for MetaPhlAn, and 0.21 for SIANN. For SIANN this means that having 38% of the species-unique genome covered by reads resulted in the vast majority of calls being accurate.
Figure 2. Relationship of reported value for each program (horizontal axis, log scale) to the empirically-determined Positive Predictive Value (PPV), shown on the vertical axis. While the exact values depend on the test data used, the general values at significant cutoff values (0.8, 0.9, 0.95 PPV) remain relatively constant across different datasets (data not shown).
For read-assignment methods (such as LMAT and Kraken), manual inspection of the results may yield a different understanding of confidence than is presented here, or in any automated analysis. For example, while each read that is assigned by LMAT and Kraken fall above a certain cutoff for species-specificity, some individual reads may be much more specific than others. One could identify a read that aligns to a single species of bacteria with 100% accuracy over its 300bp length, with the next closest match being only 90% similar. It is extremely unlikely that a 300bp exact match would arise due to random chance, and so the user could say with confidence that the organism of interest is found within the sequence data (not considering contamination, horizontal gene transfer, etc). However, such an approach is not currently implemented in an automated method, and many of the steps needed to make that assertion are performed manually by a domain expert, including alignment to near neighbors and ensuring that the read does not fall within a transposon, plasmid, etc. Therefore, while one could say that a single read is all that is needed to state with high PPV that an organism is present, the amount of reads assigned in an automated manner needed to achieve that level of PPV will number in the thousands (Fig 2).
The next phase of benchmarking was to determine how many raw input reads were needed to achieve the threshold for high PPV. To demonstrate this, we plotted the known abundance of each spike organism against the PPV value generated by each method (Figure 3). Each point (an organism at a known level of abundance) is comprised of a maximum of 50 replicates, where the diameter of each point increases with an increasing number of replicates. For demonstration purposes, we are showing two pairs of bacteria and three viruses. Recall that for each of the pairs of bacteria (and the two poxviruses) any sample containing one did not contain the other (as shown in Table 1). The empty boxes result from the organisms not being called at any abundance. For MetaPhlAn, that is a result of no viruses being included in the version of the reference database available for this analysis. Kraken assigned no reads to Hanta virus because viral RNA genomes were not included in this version of the reference database (personal communication with D. Wood). This emphasizes the point that a) the ability to create custom databases targeting organisms of interest can be valuable, and b) the performance of any method must be benchmarked against each potential target of interest.

All methods were able to identify the bulk of organisms in their databases at high abundances (75% and 18%, Figure 3), however performance varied considerably at lower abundances and depended on the particular organism and method used. SIANN detected each organism at high confidence, even at levels as low as 0.3% and 0.07% of the total.

**Discussion**

The process of detecting trace amounts of a specific organism in a complex mixture of DNA is challenging enough for an expert, but that pales in comparison to the difficulty of accomplishing the same certainty of detection in an automated manner. The results presented here show that SIANN rapidly detects the presence of a given set of organisms with a high degree of specificity and sensitivity. For example, at the 95% confidence (PPV) cutoff of 0.2, SIANN reliably detects all of the organisms tested here at as low as 0.3% abundance. This strong performance is likely due to the fact that SIANN is able to use a method (read alignment to whole genomes) that would be far too computationally costly if it were applied to the entire collection of known genomes. By focusing on a set of (customizable) target organisms and their near neighbors, SIANN can operate quickly and with low computational requirements while delivering highly accurate results.

SIANN: Strain Identification by Alignment to Near Neighbors

BaseSpace was chosen as an appropriate release platform because while the entire set of software and dependencies can be deployed by the user from within a graphical user interface, the actual computation takes place in a controlled ‘cloud’ environment. Such a distribution strategy obviates the need to satisfy the multiple software or OS dependencies that often arises with academic computational methods. Results for SIANN are compiled into a report format, showing both the organisms that surpass 95% confidence, as well as the closest strain match for each species. The default view masks the raw data output, so that the results are human-readable and do not present extraneous information. While the code for execution and database-construction on a users system is available from Signature Science, LLC, additional databases on the BaseSpace platform can be made available upon request.

There is a neverending list of questions that one could ask of metagenomic sequencing data generated from important samples. Instead of answering them all, we demonstrate a technique with a very narrow focus that is able to report with a high degree of confidence whether a given set of organisms is present in a sample. These results are presented to the user in a comprehensible format, and accessible on a commonly-used web platform. The world of bioinformatics will continue to progress and develop more sophisticated tools for metagenomic analysis, and we hope that the utility of SIANN will convince others to package and benchmark their tools in a way that they can be used with confidence by the larger public, as well as the research community.
SIANN: Strain Identification by Alignment to Near Neighbors

References


SIANN: Strain Identification by Alignment to Near Neighbors


Wood DE and Salzberg SL. Ultrafast metagenomic sequence classification using exact alignments. In submission.

SIANN: Strain Identification by Alignment to Near Neighbors

Appendices

Appendix 1: Target Pathogen Database

<table>
<thead>
<tr>
<th>Virus Family</th>
<th>Virus Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arenaviridae</td>
<td>Arenavirus</td>
</tr>
<tr>
<td></td>
<td>Golden-Gate-virus</td>
</tr>
<tr>
<td></td>
<td>Arenavirus Lujo-virus</td>
</tr>
<tr>
<td></td>
<td>Flexal-virus</td>
</tr>
<tr>
<td></td>
<td>Alphavirus</td>
</tr>
<tr>
<td></td>
<td>Machupo-virus</td>
</tr>
<tr>
<td></td>
<td>Sabia-virus</td>
</tr>
<tr>
<td></td>
<td>Tacaribe-virus</td>
</tr>
<tr>
<td></td>
<td>Whitewater-Arroyo-virus</td>
</tr>
<tr>
<td></td>
<td>New-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Allpahuayo-virus</td>
</tr>
<tr>
<td></td>
<td>Allpahuayo-virus</td>
</tr>
<tr>
<td></td>
<td>Chapare-virus</td>
</tr>
<tr>
<td></td>
<td>Guanarito-virus</td>
</tr>
<tr>
<td></td>
<td>Junin-virus</td>
</tr>
<tr>
<td></td>
<td>New-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Machupo-virus</td>
</tr>
<tr>
<td></td>
<td>Sabia-virus</td>
</tr>
<tr>
<td></td>
<td>Tacaribe-virus</td>
</tr>
<tr>
<td></td>
<td>Whitewater-Arroyo-virus</td>
</tr>
<tr>
<td></td>
<td>New-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Ippy-virus</td>
</tr>
<tr>
<td></td>
<td>Lassa-virus</td>
</tr>
<tr>
<td></td>
<td>Mopeia-virus</td>
</tr>
<tr>
<td></td>
<td>Old-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Ippy-virus</td>
</tr>
<tr>
<td></td>
<td>Old-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Old-world-arenaviruses</td>
</tr>
<tr>
<td></td>
<td>Sabia-virus</td>
</tr>
<tr>
<td></td>
<td>Tacaribe-virus</td>
</tr>
<tr>
<td></td>
<td>Whitewater-Arroyo-virus</td>
</tr>
<tr>
<td></td>
<td>Lassa-virus</td>
</tr>
<tr>
<td></td>
<td>Mopeia-virus</td>
</tr>
<tr>
<td></td>
<td>African-swine-fever-virus</td>
</tr>
<tr>
<td></td>
<td>Benin-971-pathogenic-isolate</td>
</tr>
<tr>
<td></td>
<td>African-Swine-Fever-Virus</td>
</tr>
<tr>
<td></td>
<td>Bacillus anthracis A2012 Bant021</td>
</tr>
<tr>
<td></td>
<td>Bacillus anthracis Ames Ancestor</td>
</tr>
<tr>
<td></td>
<td>Bacillus anthracis Sterne</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus 03BB102</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus AH187</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus AH820</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus ATCC 14579</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus B4264</td>
</tr>
<tr>
<td></td>
<td>Bacillus cereus F65185</td>
</tr>
<tr>
<td></td>
<td>Bacillus mycoides DSM 2048</td>
</tr>
<tr>
<td></td>
<td>Bacillus mycoides Rock1-4</td>
</tr>
<tr>
<td></td>
<td>Bacillus thuringiensis BMB171</td>
</tr>
<tr>
<td></td>
<td>Bacillus thuringiensis B1407</td>
</tr>
<tr>
<td></td>
<td>Bacillus thuringiensis HD-771</td>
</tr>
<tr>
<td></td>
<td>Bacillus thuringiensis serovar chinensis CT-43</td>
</tr>
<tr>
<td></td>
<td>Bacillus thuringiensis serovar konkukian 97-27</td>
</tr>
<tr>
<td></td>
<td>Brucella abortus A13334</td>
</tr>
<tr>
<td></td>
<td>Brucella ceti B1 94</td>
</tr>
<tr>
<td></td>
<td>Brucella ceti M13 05 1 supercont1 22</td>
</tr>
<tr>
<td></td>
<td>Brucella melitensis ATCC 23457</td>
</tr>
<tr>
<td></td>
<td>Brucella melitensis bv 1 str 16M</td>
</tr>
<tr>
<td></td>
<td>Brucella ovis ATCC 25840</td>
</tr>
<tr>
<td></td>
<td>Brucella suis 1330</td>
</tr>
<tr>
<td></td>
<td>Brucella suis ATCC 23445</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Akabane-virus segment-M</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Andes-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Dobrava-Belgrade-virus-strain-DOBV-Ano-Porea-All9-1999</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Hantaan-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Puumala-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Seoul-virus-strain-Seoul-80-39-clone-1</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus Sin-Nombre-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus</td>
</tr>
<tr>
<td></td>
<td>Crimean-Congo-hemorrhagic-fever-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Hantavirus Tula-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Nairobi virus</td>
</tr>
<tr>
<td></td>
<td>Dugbe-virus</td>
</tr>
<tr>
<td></td>
<td>Bunyaviridae Phlebovirus Rift-Valley-fever-virus</td>
</tr>
<tr>
<td></td>
<td>Burkholderia cepacia GG4</td>
</tr>
<tr>
<td></td>
<td>Burkholderia gladioli BSR3</td>
</tr>
<tr>
<td></td>
<td>Burkholderia gluma BGR1</td>
</tr>
<tr>
<td></td>
<td>Burkholderia mallei</td>
</tr>
<tr>
<td></td>
<td>Burkholderia mallei ATCC 10399</td>
</tr>
<tr>
<td></td>
<td>Burkholderia mallei NCTC 10247</td>
</tr>
<tr>
<td></td>
<td>Burkholderia mallei SAVP1</td>
</tr>
<tr>
<td></td>
<td>Burkholderia multivorans</td>
</tr>
<tr>
<td></td>
<td>Burkholderia oklahomensis  C6786</td>
</tr>
<tr>
<td></td>
<td>Burkholderia oklahomensis  EO147</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  1026b</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  1106a</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  1710b</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  668</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  BPC006</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pseudomallei  K96243</td>
</tr>
<tr>
<td></td>
<td>Burkholderia pyrocinia CH-67</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis ATCC 700388</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis E264</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis MSMB121</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis Campylobacter coli JV20</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis Campylobacter fetus subsp fetus 82-40</td>
</tr>
<tr>
<td></td>
<td>Burkholderia thailandensis Campylobacter jejuni RM1221</td>
</tr>
</tbody>
</table>
Campylobacter jejuni subsp. doylei 26997
Campylobacter jejuni subsp. jejuni B1-176
Campylobacter jejuni subsp. jejuni NCTC 11168 ATCC 700819
Campylobacter upsaliensis JV21
Clostridium acetobutylicum DSM 1731
Clostridium botulinum A str. ATCC 3502
Clostridium botulinum BKT015925
Clostridium botulinum B str. Eklund 17B
Clostridium botulinum E3 str. Alaska E43
Clostridium botulinum F str. 230613
Clostridium botulinum H04402 065
Clostridium difficile 2007855
Clostridium difficile 630
Clostridium difficile BI1
Clostridium difficile BI9
Clostridium perfringens ATCC 13124
Clostridium perfringens SM101
Clostridium perfringens str 13
Clostridium symbiosum WAL-14163
Clostridium symbiosum WAL-14673
Clostridium tetani E88
Clostridium thermocellum ATCC 27405
Clostridium thermocellum DSM 1313
Clostridium tunisiense TJ C661
Clostridium ultunense Esp
Coccidioides immitis H5384
Coccidioides immitis RMSCC 2394
Coccidioides immitis RS
Coccidioides posadasi C735
delta SOWgp
Coccidioides posadasi CPA 0001
Coccidioides posadasi CPA 0020
Coronaviridae
Alphacoronavirus Bat-coronavirus-HKU2
Coronaviridae
Alphacoronavirus Feline-infectious-peritonitis-virus
Coronaviridae
Alphacoronavirus Human-coronavirus-229E
Coronaviridae
Alphacoronavirus TGEV-Purdue-P115
Coronaviridae Bafinivirus White-bream-virus
Coronaviridae Betacoronavirus Bovine-coronavirus
Coronaviridae Betacoronavirus Murine-hepatitis-virus-strain-A59
Coronaviridae Betacoronavirus Murine-hepatitis-virus-strain-JHM
Coronaviridae Betacoronavirus SARS-coronavirus
Coronaviridae Coronavirinae Munia-coronavirus-HKU13-3514
Coronaviridae Gammacoronavirus Avian-infectious-bronchitis-virus
Coronaviridae Gammacoronavirus Turkey-coronavirus
Coronaviridae Torovirus Breda-virus
Coronaviridae unclassified-coronaviruses Bat-coronavirus-BM48-31-BGR-2008
Coronaviridae unclassified-coronaviruses Bovine-respiratory-coronavirus-AH187
Coronaviridae unclassified-coronaviruses Human-enteric-coronavirus-strain-4408
Coxiella burnetii CbuG Q212
Coxiella burnetii Dugway 5J108 111
Coxiella burnetii RSA 493
Diplococcietsis massiliensis 208 CS 1
Ehrlichia canis str Jake
Ehrlichia chaffeensis str. Arkansas
Ehrlichia chaffeensis str. Sapulpa ctg90
Ehrlichia ruminantium str. Gardel
Ehrlichia ruminantium str. Welgevonden
Escherichia coli APEC O1
Escherichia coli BL21 DE3
Escherichia coli B str REL606
Escherichia coli ETEC H10407
Escherichia coli O104H4 str. 2011C-3493
Escherichia coli O157H7 str. EC4115
Escherichia coli O157H7 str. Sakai
Escherichia coli O7K1 str CE10
Escherichia coli O83H1 str. NRG 857C
Escherichia coli str. K-12 substr. MG1655
Filoviridae Ebolavirus Bundibugyo-ebolavirus
Filoviridae Ebolavirus Cote-d'Ivoire-ebolavirus
Filoviridae Ebolavirus Ebola-virus-Mayinga-Zaire
Filoviridae Ebolavirus Reston-ebolavirus
Filoviridae Ebolavirus Sudan-ebolavirus
Filoviridae Marburgvirus Lake-Victoria-marburgvirus-Musoke
Flaviviridae Alkhurma-virus
Flaviviridae Classical-swine-fever-virus
Flaviviridae Dengue-virus 1
Flaviviridae Dengue-virus 2
Flaviviridae Dengue-virus 3
Flaviviridae Dengue-virus 4
Flaviviridae Japanese-encephalitis-virus genome
Flaviviridae Karshi-virus
<table>
<thead>
<tr>
<th>Family</th>
<th>Virus Type</th>
<th>Species/Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flaviviridae</td>
<td>Langat-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Louping-ill-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Murray-Valley-encephalitis-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Omsk-hemorrhagic-fever-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Powassan-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>St-Louis-encephalitis-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Tick-borne-encephalitis-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>Usutu-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Flaviviridae</td>
<td>West-Nile-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Herpesviridae</td>
<td>Alcelaphine-herpesvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Herpesviridae</td>
<td>Macacine-herpesvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>novicida Fx1</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>novatensis str Toba 04</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>philomiragia subsp philomiragia ATCC 25015</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>philomiragia subsp philomiragia ATCC 25017</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>tularensis subsp holarctica F92</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Francisella</td>
<td>tularensis subsp mediasiatica FSC147</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Herpesviridae</td>
<td>Acelaphine-herpesvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Herpesviridae</td>
<td>Macacine-herpesvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>fleischmannii LU2006-1 c88</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>innocua Clip11262</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>ivanovii FSL J1-023</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>ivanovii FSL 07PF0776</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes 08-5578</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes 10403S</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes ATCC 19117</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes Finland 1998</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes FSL R2-561</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Listeria</td>
<td>monocytogenes La111</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Lumpy-skin-disease-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Molluscipoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Molluscum-contagiosum-virus-subtype-1</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Camelpox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Cowpox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>ECTROMELIA-VIRUS</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Monkeypox-virus-Zaire-96-1-16</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Taterapox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Vaccinia-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxivirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Sheepox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Suipoxvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Swinepox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Puccinia graminis f sp tritici</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Camelpox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Cowpox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Ectromelia-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Swainula-spike-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Peplit-sue-spiral-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Rinderpest-virus-strain-Kabete-O</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Rubulavirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Human-parainfluenza-virus-1</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Human-parainfluenza-virus-2</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Human-parainfluenza-virus-4-a</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Rubulavirus-Mumps-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Foot-and-mouth-disease-virus-type-O</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Swine-vesicular-disease-virus-strain-HK70</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Swine-vesicular-disease-virus-strain-NET192</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Avipoxvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Fowlipox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Crocodilepoxvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Nile-crocodilepox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Goatpox-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Pellor</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Leporipoxvirus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
<tr>
<td>Poxviridae</td>
<td>Orthopoxvirus-Myxoma-virus</td>
<td>SIANN: Strain Identification by Alignment to Near Neighbors</td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Salmonella bongori N268-08
Salmonella bongori NCTC 12419
Salmonella enterica subsp arizonae serovar 62z4z23- str RSK2980
Salmonella enterica subsp enterica serovar Dublin str CT 02021853
Salmonella enterica subsp enterica serovar Newport str SL254
Salmonella enterica subsp enterica serovar Paratyphi A str ATCC 9150
Salmonella enterica subsp enterica serovar Typhimurium str LT2
Salmonella enterica subsp enterica serovar Typhi str Ty2
Shigella boydii CDC 3083-94
Shigella boydii Sb227
Shigella dysenteriae Sd197
Shigella flexneri 2002017
Shigella flexneri 2a str 2457T
Shigella flexneri 2a str 301
Shigella flexneri 5 str 8401
Shigella sonnei 53G
Shigella sonnei Ss046
Staphylococcus arlettae CVD059 SARL c230
Staphylococcus aureus 04-02981
Staphylococcus aureus 08BA02176
Staphylococcus aureus subsp aureus N315
Staphylococcus aureus subsp aureus NCTC 8325
Staphylococcus aureus subsp aureus TW20
Staphylococcus capitis QN1 Contig63
Staphylococcus capitis SK14
Staphylococcus caprae C87
Staphylococcus carnosus subsp carnosus TM300
Staphylococcus epidermidis ATCC 12228
Staphylococcus epidermidis RP62A
Staphylococcus equorum subsp equorum Mu2
Staphylococcus haemolyticus JCSC1435
Staphylococcus hominis SK119
Staphylococcus hominis subsp hominis C80
Staphylococcus lugdunensis HKU09-01
Staphylococcus lugdunensis N920143
Togaviridae Alphavirus Barmah-Forest-virus
Togaviridae Chikungunya-virus
Togaviridae EEEV-complex Eastern-equine-encephalitis-virus
Togaviridae Rubivirus Rubella-virus
Togaviridae SFV-complex O-nyong-nyong-virus
Togaviridae Venezuealan-equine-encephalitis-virus
Togaviridae WEEV-complex Sindbis-virus
Togaviridae Western-equine-encephalomyelitis-virus
Xanthomonas albilineans GPE PC73
Xanthomonas axonopodi Xac29-1
Xanthomonas oryzae pv oryzae KACC 10331
Xanthomonas oryzae pv oryzae MAFF 311018
Xanthomonas oryzae pv oryzae PXO99A
Xanthomonas vasicola pv vascularum NCPPB 1326 scf 9767 4580
Yersinia aldovae ATCC 35236
Yersinia bercovieri ATCC 43970
Yersinia enterocolitica IP 10393
Yersinia enterocolitica IP2222
Yersinia enterocolitica subsp enterocolitica 8081
Yersinia enterocolitica subsp palearctica 105 SR
Yersinia frederiksenii ATCC 33641
Yersinia intermedia ATCC 29909
Yersinia kristensenii ATCC 33638
Yersinia mollaretii ATCC 43969
Yersinia pestis A1122
Yersinia pestis Antiqua
Yersinia pestis KIM 10
Yersinia pestis Pestoides F
Yersinia pseudotuberculosis IP 31758
Yersinia pseudotuberculosis IP 32953
Yersinia pseudotuberculosis PB1
Yersinia pseudotuberculosis YPIII
Yersinia rohdei ATCC 43380
Yersinia ruckeri ATCC 29473
Appendix 2: Viral Database

Abaca bunchy top virus DNA-C
Abaca bunchy top virus DNA-M
Abaca bunchy top virus DNA-N
Abaca bunchy top virus DNA-R
Abaca bunchy top virus DNA-S
Abaca bunchy top virus segment 2
Abalone shriveling syndrome-associated virus
Abutilon Brazil virus DNA A
Abutilon Brazil virus DNA B
Abutilon mosaic virus DNA A
Abutilon mosaic virus DNA B
Acanthocystis turfacea
Chlorella virus 1
Acheta domesticus
densovirus
Acholeplasma phage L2
Acholeplasma phage MV-L1
Acidianus bottle-shaped virus
Acidianus filamentous virus 1
Acidianus filamentous virus 2
Acidianus filamentous virus 3
Acidianus filamentous virus 6
Acidianus filamentous virus 7
Acidianus filamentous virus 8
Acidianus filamentous virus 9
Acidianus rod-shaped virus 1
Acidianus spindle-shaped virus 1
Acidianus two-tailed virus
Actinomycyes phage Av-1
Actinoplane phage phiAsp2
Acyrtosiphon pisum
bacteriophage APSE-1
Adeno-associated virus-1
Adeno-associated virus-2
Adeno-associated virus-3
Adeno-associated virus-4
Adeno-associated virus-5
Adeno-associated virus-7
Adeno-associated virus-8
Adoxophyes honmai NPV

Adoxophyes orana
granulovirus
Adoxophyes orana
nucleopolyhedrovirus
Aedes aegypti
densovirus
Aedes albopictus
densovirus
Aedes taeniorynchus
iridescent virus
Aeromonas phage 25
Aeromonas phage 31
Aeromonas phage 44RR2.8t
Aeromonas phage phiO18P
african cassava mosaic virus
DNA A
african cassava mosaic virus
DNA B
African green monkey
polyomavirus
African swine fever virus
Ageratum enation virus
Ageratum leaf Cameroon
betasatellite
Ageratum leaf curl virus-G52
Ageratum yellow vein China
virus-associated DNA beta
Ageratum yellow vein
Chinavirus
Ageratum yellow vein Hualian
virus-TaiwanHsinchutom2003
DNA A
Ageratum yellow vein Sri
Lanka virus segment A
Ageratum yellow vein Taiwan
virus
Ageratum yellow vein
virusassociated DNA beta
Ageratum yellow vein virus
multiple
nucleopolyhedrovirus
Agrotis segetum
granulovirus
Agrotis segetum
nucleopolyhedrovirus
Alcelaphine herpesvirus 1
Aleutian mink disease virus
Allamanda leaf curl virus
DNA-A
Alternanthera yellow vein
virus satellite DNA beta
Ambystoma tigrinum virus
Amsacta moorei
entomopoxivirus L
Anguillid herpesvirus 1
Anopheles gambiae
densovirus
Anticersia gennatals
nucleopolyhedrovirus
Archeal BJ1 virus
Ateline herpesvirus 3
Autographa californica
cassava mosaic virus
Avian adeno-associated virus
ATCC VR-865
Avian adeno-associated virus
strain DA-1
Avian endogenous retrovirus
EAV-HP
Azospirillum phage Cd
Bacillus phage 0305phi8-36
Bacillus phage AP50
Bacillus phage B103
Bacillus phage Bam35c
Bacillus phage BCJA1c
Bacillus phage Cherry
Bacillus phage Fah
Bacillus phage GA-1
Bacillus phage Gamma
Bacillus phage GIL16c
Bacillus phage IEBH
Bacillus phage phi105
Bacillus phage phi29
Bacillus phage SPBc2
Bacillus phage SPO1
Bacillus phage SPP1
Bacillus phage TP21-L
Bacillus phage WBeta
Bacillus prophage phBC6A51
Bacillus prophage phBC6A52
Bacillus virus 1
Bacteriophage Aaphi23
Bacteriophage APSE-2
Bacteriophage PSA
Bacteriophage RB32
Bacteroides phage B40-8
SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Banana bunchy top virus DNA C</th>
<th>Begomovirus-associated DNA II</th>
<th>Burkholderia ambifaria phage BcepF1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banana bunchy top virus DNA M</td>
<td>Begomovirus-associated DNA-III</td>
<td>Burkholderia phage Bcep1</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA N</td>
<td>Bettongia penicillata papillomavirus 1</td>
<td>Burkholderia phage Bcep22</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA R</td>
<td>Bhendi yellow vein virus DNA-A</td>
<td>Burkholderia phage Bcep43</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA S</td>
<td>Bhendi yellow vein Delhi virus 2004New Delhi DNA-A</td>
<td>Burkholderia phage Bcep781</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA U3</td>
<td>Bhendi yellow vein mosaic virus-associated DNA beta</td>
<td>Burkholderia phage Bcep81A</td>
</tr>
<tr>
<td>Banana streak GF virus</td>
<td>Bhendi yellow vein mosaic Virus</td>
<td>Burkholderia phage BcepC68</td>
</tr>
<tr>
<td>Banana streak Mysore virus</td>
<td>Bitter gourd leaf curl disease-associated DNA beta</td>
<td>Burkholderia phage BcepGoMr</td>
</tr>
<tr>
<td>Banana streak OL virus</td>
<td>BK polyomavirus</td>
<td>Burkholderia phage BcepIL02</td>
</tr>
<tr>
<td>Banana streak virus genome</td>
<td>Blainvillaea yellow spot virus DNA-A</td>
<td>Burkholderia phage BcepMu</td>
</tr>
<tr>
<td>Banana streak virus strain Acuminata Vietnam</td>
<td>Blainvillaea yellow spot virus DNA-B</td>
<td>Burkholderia phage BcepNazgul</td>
</tr>
<tr>
<td>Bandicoot papillomatosis carcinoma virus type 1</td>
<td>Blattella germanica densovirus</td>
<td>Burkholderia phage BcepNY3</td>
</tr>
<tr>
<td>Bandicoot papillomatosis carcinoma virus type 2</td>
<td>Blueberry red ringspot virus</td>
<td>Burkholderia phage KS10</td>
</tr>
<tr>
<td>Bat adeno-associated virus YNM</td>
<td>Bocavirus gorillaGBov12009</td>
<td>Burkholderia phage KS9</td>
</tr>
<tr>
<td>Bdellovibrio phage phiMH2K</td>
<td>Bombyx mandarina nucleopolyhedrovirus</td>
<td>Burkholderia phage phi1026b</td>
</tr>
<tr>
<td>Beak and feather disease virus</td>
<td>Bombyx mori densovirus 5</td>
<td>Burkholderia phage phi644-2 chromosome</td>
</tr>
<tr>
<td>Bean calico mosaic virus DNA A</td>
<td>Bombyx mori NPV</td>
<td>Burkholderia phage phiE12-2 chromosome</td>
</tr>
<tr>
<td>Bean calico mosaic virus DNA B</td>
<td>Bordetella phage BIP-1</td>
<td>Burkholderia phage phiE125</td>
</tr>
<tr>
<td>Bean dwarf mosaic virus DNA A</td>
<td>Bordetella phage BMP-1</td>
<td>Burkholderia phage phiE202 chromosome</td>
</tr>
<tr>
<td>Bean dwarf mosaic virus DNA B</td>
<td>Bordetella phage BPP-1</td>
<td>Burkholderia phage phiE255 chromosomeBurkholderia prophage phi52237</td>
</tr>
<tr>
<td>Bean golden mosaic virus DNA A</td>
<td>Bougainvillaea spectabilis chlorotic vein-banding virus</td>
<td>Cabbage leaf curl virus DNA A</td>
</tr>
<tr>
<td>Bean golden mosaic virus DNA B</td>
<td>Bovine aden-associated virus</td>
<td>Cabbage leaf curl virus DNA B</td>
</tr>
<tr>
<td>Bean golden yellow mosaic virus DNA A</td>
<td>Bovine adenovirus A</td>
<td>Cacao swollen shoot virus</td>
</tr>
<tr>
<td>Bean golden yellow mosaic virus DNA B</td>
<td>Bovine adenovirus B</td>
<td>California sea lion anellovirus</td>
</tr>
<tr>
<td>Bean yellow dwarf virus putative genes V1</td>
<td>Bovine adenovirus D</td>
<td>California sea lion polyomavirus 1</td>
</tr>
<tr>
<td>Beet curly top Iran virus-K</td>
<td>Bovine ephemeral fever virus</td>
<td>Callitrichine herpesvirus 3 strain CJ0149</td>
</tr>
<tr>
<td>Beet curly top virus-California Logan</td>
<td>Bovine foamy virus</td>
<td>Camelpox virus</td>
</tr>
<tr>
<td>Beet mild curly top virus-Worland4</td>
<td>Bovine herpesvirus 1</td>
<td>Campoletis sonorenensis ichnovirus chromosome segment W</td>
</tr>
<tr>
<td>Beet severe curly top virus-Cfh</td>
<td>Bovine herpesvirus 4 long unique region</td>
<td>Campoletis sonorenensis ichnovirus segment B</td>
</tr>
<tr>
<td></td>
<td>Bovine herpesvirus 5</td>
<td>Campoletis sonorenensis ichnovirus segment C</td>
</tr>
<tr>
<td></td>
<td>Bovine papillomavirus-1</td>
<td>Campoletis sonorenensis ichnovirus segment D</td>
</tr>
<tr>
<td></td>
<td>Bovine papillomavirus 3</td>
<td>Campoletis sonorenensis ichnovirus segment E</td>
</tr>
<tr>
<td></td>
<td>Bovine papular stomatitis virus</td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Virus Name</th>
<th>Segment/Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>F</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>G2</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>G</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>H</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>I2</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>I</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>J</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>L</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>M</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>N</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>O1</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>P</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>Q</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>T</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>U</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>V</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus</td>
<td>Z</td>
</tr>
<tr>
<td>Campoletis sonorensis ichnovirus superhelical</td>
<td></td>
</tr>
<tr>
<td>Virus Name</td>
<td>Segment/Region</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle12</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle13</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle14</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle15</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle17</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle19</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment circle1</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle20</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle21</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle22</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle23</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle25</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle26</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment circle2</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle30</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle31</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle32</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle33</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle35</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle36</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment circle3</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle4</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle5</td>
</tr>
<tr>
<td>Cotesia congregata</td>
<td>bracovirus segment Circle9</td>
</tr>
<tr>
<td>Cotesia congregata virus stage Circle18</td>
<td></td>
</tr>
<tr>
<td>Cotesia congregata virus stage Circle6</td>
<td></td>
</tr>
<tr>
<td>Cotesia congregata virus stage Circle7</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf crumple geminivirus DNA B</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf crumple virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Alabad virus</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Bangalore virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Bangalore virus segment A</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Burewala alphasatellite</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Burewala betasatellite</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Burewala virus-IndiaVehari2004</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Gezira alphasatellite</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Gezira beta</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Gezira Betasatellite</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Gezira extrachromosomal</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Gezira virus</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Kokhran virus</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Multan Virus</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Multan virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Multan virus satellite U36-1</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl Rajasthan virus segment A</td>
<td></td>
</tr>
<tr>
<td>Cotton leaf curl virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Cottontail rabbit papillomavirus</td>
<td></td>
</tr>
<tr>
<td>Cowpea severe leaf curl-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Cowpox virus</td>
<td></td>
</tr>
<tr>
<td>Crassocephalum yellow vein virus-Jinghong</td>
<td></td>
</tr>
<tr>
<td>Croton yellow vein mosaic alphasatellite</td>
<td></td>
</tr>
<tr>
<td>Croton yellow vein mosaic virus</td>
<td></td>
</tr>
<tr>
<td>Croton yellow vein mosaic Virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Croton yellow vein virus</td>
<td></td>
</tr>
<tr>
<td>Crow polyomavirus</td>
<td></td>
</tr>
<tr>
<td>Cryptophlebia leucotreta granulovirus</td>
<td></td>
</tr>
<tr>
<td>Cucurbita yellow vein virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Cucurbita leaf crumple virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Cucurbita leaf crumple virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Culex nigripalpus NPV</td>
<td></td>
</tr>
<tr>
<td>Culex pipiens densovirus</td>
<td></td>
</tr>
<tr>
<td>Cyanophage P52</td>
<td></td>
</tr>
<tr>
<td>Cyanophage Syn5</td>
<td></td>
</tr>
<tr>
<td>Cycad leaf necrosis virus</td>
<td></td>
</tr>
<tr>
<td>Cydia pomonella granulovirus</td>
<td></td>
</tr>
<tr>
<td>Cyprinid herpesvirus 3</td>
<td></td>
</tr>
<tr>
<td>Deer papillomavirus</td>
<td></td>
</tr>
<tr>
<td>Deerpox virus W-1170-84</td>
<td></td>
</tr>
<tr>
<td>Deerpox virus W-848-83</td>
<td></td>
</tr>
<tr>
<td>Deftia phage phiW-14</td>
<td></td>
</tr>
<tr>
<td>Dendrolimus punctatus densovirus</td>
<td></td>
</tr>
<tr>
<td>Desmodium leaf distortion virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Desmodium leaf distortion virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Diadromus pulchellus</td>
<td></td>
</tr>
<tr>
<td>ascovirus 4a</td>
<td></td>
</tr>
<tr>
<td>Diatraea saccharalis densovirus</td>
<td></td>
</tr>
<tr>
<td>Diciplerta yellow mottle virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Diciplerta yellow mottle virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Digitaria streak virus</td>
<td></td>
</tr>
<tr>
<td>Dioscorea bacilliform virus</td>
<td></td>
</tr>
<tr>
<td>Dolichos yellow mosaic virus</td>
<td></td>
</tr>
<tr>
<td>Dracaena mottle virus</td>
<td></td>
</tr>
<tr>
<td>Duck adenovirus A</td>
<td></td>
</tr>
<tr>
<td>Duck circovirus</td>
<td></td>
</tr>
<tr>
<td>Duck hepatitis B virus</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Cameroon virus DNA A</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Cameroon virus DNA</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic virus</td>
<td></td>
</tr>
<tr>
<td>Kenya virus DNA A</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic virus</td>
<td></td>
</tr>
<tr>
<td>Kenya virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Virus/Phage/Species</td>
<td>Type</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>East African cassava mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic Zanzibar virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>East African cassava mosaic Zanzibar virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Ecotropis obliqua NPV</td>
<td></td>
</tr>
<tr>
<td>Ectocarpus siliculosus virus 1</td>
<td></td>
</tr>
<tr>
<td>Ectromelia virus</td>
<td></td>
</tr>
<tr>
<td>Emiliana huxleyi virus 86</td>
<td></td>
</tr>
<tr>
<td>Emilia yellow vein virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage 13a</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage 933W</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage alpha3</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage 8A14</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage BP-4795</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage EcoDS1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage EPS7</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage epsilon15</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage ES18</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Felix 01</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Fels-2</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage G4</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage HK022</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage HK620</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage HK97</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage I2-2</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage ID18</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage ID2</td>
<td></td>
</tr>
<tr>
<td>MoscowI2001</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage If1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Ike</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage IME08</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage JK06</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage JS10</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage JS98</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage JSE</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage K1-5</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage K1E</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage K1F</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage lambda</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage LKA1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage M13</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Min27</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Mu</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage N15</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage N4</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage P1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage P22 virus</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage P2 virus</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage P4</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage Phi</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage phiEco32</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage phiEcoM-GJ1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage phiP27</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage phiX174 sensu lato</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage PRD1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage PsP3</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB14</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB16</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB43</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB49</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB51</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RB69</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage RTP</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage S16</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage SIV</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage SP6</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage SSL-2009a</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage ST104</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage S1-1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage ST64T</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage T1</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage T3</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage T4</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage T5</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage T7</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage TLS</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage VT2-Sakai</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage WA13</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage WV8</td>
<td></td>
</tr>
<tr>
<td>Enterobacteria phage YYZ-2008</td>
<td></td>
</tr>
<tr>
<td>Enterococcus faecalis 62 plasmid EF62pB</td>
<td></td>
</tr>
<tr>
<td>Enterococcus faecalis 62 plasmid EF62pC</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage EF62phi</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage EFAP-1</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiEF11</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiEF24C</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiFL1A</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiFL2A</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiFL3A</td>
<td></td>
</tr>
<tr>
<td>Enterococcus phage phiFL4A</td>
<td></td>
</tr>
<tr>
<td>Enzootic nasal tumour virus of goats</td>
<td></td>
</tr>
<tr>
<td>Epiphysa postvittana NPV</td>
<td></td>
</tr>
<tr>
<td>Equid herpesvirus 1</td>
<td></td>
</tr>
<tr>
<td>Equid herpesvirus 2</td>
<td></td>
</tr>
<tr>
<td>Equid herpesvirus 4</td>
<td></td>
</tr>
<tr>
<td>Equid herpesvirus 9</td>
<td></td>
</tr>
<tr>
<td>Equine papillomavirus 2</td>
<td></td>
</tr>
<tr>
<td>Eragrostis curvula streak virus</td>
<td></td>
</tr>
<tr>
<td>Eragrostis streak virus</td>
<td></td>
</tr>
<tr>
<td>Erectites yellow mosaics virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Erectites yellow mosaics virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Erethizon dorsatum papillomavirus type 1</td>
<td></td>
</tr>
<tr>
<td>Erinaceus europaeus papillomavirus</td>
<td></td>
</tr>
<tr>
<td>Erwinia amylovora phage Era103</td>
<td></td>
</tr>
<tr>
<td>Erwinia phage phiEa21-4</td>
<td></td>
</tr>
<tr>
<td>Escherichia coli bacteriophage rv5</td>
<td></td>
</tr>
<tr>
<td>Escherichia phage D108</td>
<td></td>
</tr>
<tr>
<td>Escherichia phage phiV10</td>
<td></td>
</tr>
<tr>
<td>Eupatorium vein clearing virus</td>
<td></td>
</tr>
<tr>
<td>Eupatorium yellow vein associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Eupatorium yellow vein virus</td>
<td></td>
</tr>
<tr>
<td>Euphorbia leaf curl virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Euphorbia yellow mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Euphorbia yellow mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Euproctis pseudoconspersa nucleopolyhedrovirus</td>
<td></td>
</tr>
<tr>
<td>European elk papillomavirus</td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Faba bean necrotic stunt virus DNA C
Faba bean necrotic stunt virus DNA M
Faba bean necrotic stunt virus DNA N
Faba bean necrotic stunt virus DNA R
Faba bean necrotic stunt virus DNA S
Faba bean necrotic stunt virus DNA U1
Faba bean necrotic stunt virus DNA U2
Faba bean necrotic stunt virus DNA U4
Faba bean necrotic yellows virus DNA 10
Faba bean necrotic yellows virus DNA-1
Faba bean necrotic yellows virus DNA 2
Faba bean necrotic yellows virus DNA 4
Faba bean necrotic yellows virus DNA 5
Faba bean necrotic yellows virus DNA 7
Faba bean necrotic yellows virus DNA 8
Faba bean necrotic yellows virus DNA 9
Faba bean necrotic yellows Virus
Feldmannia species virus
Felis domesticus papillomavirus type 1
Fenneropenaeus chinensis hepatopancreatic densovirus
Figwort mosaic virus
Fringilla coelebs papillomavirus
Francolinus leucoscepus papillomavirus
Gallia mellonella densovirus
Gallid herpesvirus 1
Gallid herpesvirus 2
Gallid herpesvirus 3
Gammapapillomavirus
HPV127
Geobacillus phage GBSV1
Geobacillus virus E2
Glossina pallidipes salivary gland hypertrophy virus
Glypta fumiferanae ichnovirus segment A10
Glypta fumiferanae ichnovirus segment A1
Glypta fumiferanae ichnovirus segment A2
Glypta fumiferanae ichnovirus segment A3
Glypta fumiferanae ichnovirus segment A4
Glypta fumiferanae ichnovirus segment A5
Glypta fumiferanae ichnovirus segment A6
Glypta fumiferanae ichnovirus segment A7
Glypta fumiferanae ichnovirus segment A8
Glypta fumiferanae ichnovirus segment A9
Glypta fumiferanae ichnovirus segment B10
Glypta fumiferanae ichnovirus segment B11
Glypta fumiferanae ichnovirus segment B12
Glypta fumiferanae ichnovirus segment B13
Glypta fumiferanae ichnovirus segment B14
Glypta fumiferanae ichnovirus segment B15
Glypta fumiferanae ichnovirus segment B16
Glypta fumiferanae ichnovirus segment B17
Glypta fumiferanae ichnovirus segment B18
Glypta fumiferanae ichnovirus segment B19
Glypta fumiferanae ichnovirus segment B20
Glypta fumiferanae ichnovirus segment B21
Glypta fumiferanae ichnovirus segment B22
Glypta fumiferanae ichnovirus segment B23
Glypta fumiferanae ichnovirus segment B24
Glypta fumiferanae ichnovirus segment B25
Glypta fumiferanae ichnovirus segment B26
Glypta fumiferanae ichnovirus segment B27
Glypta fumiferanae ichnovirus segment B28
Glypta fumiferanae ichnovirus segment B29
Glypta fumiferanae ichnovirus segment B30
Glypta fumiferanae ichnovirus segment B31
Glypta fumiferanae ichnovirus segment B32
Glypta fumiferanae ichnovirus segment B33
Glypta fumiferanae ichnovirus segment B34
Glypta fumiferanae ichnovirus segment B35
Glypta fumiferanae ichnovirus segment B36
Glypta fumiferanae ichnovirus segment B37
Glypta fumiferanae ichnovirus segment B38
Glypta fumiferanae ichnovirus segment B39
Glypta fumiferanae ichnovirus segment B40

23
<table>
<thead>
<tr>
<th>Glypta fumiferanae ichnovirus segment B41</th>
<th>Glypta fumiferanae ichnovirus segment B42</th>
<th>Glypta fumiferanae ichnovirus segment B43</th>
<th>Glypta fumiferanae ichnovirus segment B44</th>
<th>Glypta fumiferanae ichnovirus segment B45</th>
<th>Glypta fumiferanae ichnovirus segment B46</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glypta fumiferanae ichnovirus segment B47</td>
<td>Glypta fumiferanae ichnovirus segment B48</td>
<td>Glypta fumiferanae ichnovirus segment B49</td>
<td>Glypta fumiferanae ichnovirus segment B50</td>
<td>Glypta fumiferanae ichnovirus segment B51</td>
<td>Glypta fumiferanae ichnovirus segment B52</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B53</td>
<td>Glypta fumiferanae ichnovirus segment B54</td>
<td>Glypta fumiferanae ichnovirus segment B55</td>
<td>Glypta fumiferanae ichnovirus segment B56</td>
<td>Glypta fumiferanae ichnovirus segment B57</td>
<td>Glypta fumiferanae ichnovirus segment B58</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B59</td>
<td>Glypta fumiferanae ichnovirus segment B60</td>
<td>Glypta fumiferanae ichnovirus segment B61</td>
<td>Glypta fumiferanae ichnovirus segment B62</td>
<td>Glypta fumiferanae ichnovirus segment B63</td>
<td>Glypta fumiferanae ichnovirus segment B64</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B65</td>
<td>Glypta fumiferanae ichnovirus segment B66</td>
<td>Glypta fumiferanae ichnovirus segment B67</td>
<td>Glypta fumiferanae ichnovirus segment B68</td>
<td>Glypta fumiferanae ichnovirus segment B69</td>
<td>Glypta fumiferanae ichnovirus segment B70</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B71</td>
<td>Glypta fumiferanae ichnovirus segment B72</td>
<td>Glypta fumiferanae ichnovirus segment B73</td>
<td>Glypta fumiferanae ichnovirus segment B74</td>
<td>Glypta fumiferanae ichnovirus segment B75</td>
<td>Glypta fumiferanae ichnovirus segment B76</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B77</td>
<td>Glypta fumiferanae ichnovirus segment B78</td>
<td>Glypta fumiferanae ichnovirus segment B79</td>
<td>Glypta fumiferanae ichnovirus segment B80</td>
<td>Glypta fumiferanae ichnovirus segment B81</td>
<td>Glypta fumiferanae ichnovirus segment B82</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B83</td>
<td>Glypta fumiferanae ichnovirus segment B84</td>
<td>Glypta fumiferanae ichnovirus segment B85</td>
<td>Glypta fumiferanae ichnovirus segment B86</td>
<td>Glypta fumiferanae ichnovirus segment B87</td>
<td>Glypta fumiferanae ichnovirus segment B88</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B89</td>
<td>Glypta fumiferanae ichnovirus segment B90</td>
<td>Glypta fumiferanae ichnovirus segment B91</td>
<td>Glypta fumiferanae ichnovirus segment B92</td>
<td>Glypta fumiferanae ichnovirus segment B93</td>
<td>Glypta fumiferanae ichnovirus segment B94</td>
</tr>
<tr>
<td>Glypta fumiferanae ichnovirus segment B95</td>
<td>Glypta fumiferanae ichnovirus segment B96</td>
<td>Glypta fumiferanae ichnovirus segment B97</td>
<td>Glypta fumiferanae ichnovirus segment B98</td>
<td>Glypta fumiferanae ichnovirus segment B99</td>
<td>Glypta fumiferanae ichnovirus segment B100</td>
</tr>
</tbody>
</table>

Note: The list continues with segments B101 to B300.
Halorubrum phage HF2
Halorubrum pleomorphic virus
Halovirus HF1
Hamster polyomavirus
Helicoverpa armigera granulovirus
Helicoverpa armigera multiple nucleopolyhedrovirus
Helicoverpa armigera-NPV
Helicoverpa armigera NPV NNg1
Helicoverpa armigera nucleopolyhedrovirus G4
Helicoverpa zea SNPV
Heliothis virescens ascovirus 3e
Hepatitis B virus
Heron hepatitis B virus
His1 virus
His2 virus
Hollyhock leaf crumple virus
Honeysuckle yellow vein mosaic virus bet-JapanFukui2001
Honeysuckle yellow vein mosaic virus beta-JapanMiyizaki2001
Honeysuckle yellow vein mosaic disease associated satellite DNA beta-Ibaraki
Honeysuckle yellow vein mosaic Virus
Honeysuckle yellow vein mosaic virus-Kagoshima
Honeysuckle yellow vein mosaic virus satellite DNA beta
Honeysuckle yellow vein virus-UK1
Horsegram yellow mosaic Virus DNA B
Horsegram yellow mosaic virus
Horseradish curly top virus
Human adenovirus 54
Human adenovirus A
Human adenovirus B1
Human adenovirus B2
Human adenovirus C
Human adenovirus D
Human adenovirus E
Human adenovirus F
Human bocavirus 1
Human bocavirus 2
Human bocavirus 3
Human bocavirus 4
Human bocavirus V9
Human herpesvirus 1
Human herpesvirus 2
Human herpesvirus 3
Human herpesvirus 4
Human herpesvirus 5 strain Merlin
Human herpesvirus 7
Human herpesvirus 8
Human papillomavirus-18
Human papillomavirus 1
Human papillomavirus-2
Human papillomavirus 54
Human papillomavirus-5
Human papillomavirus type-101
Human papillomavirus type-103
Human papillomavirus type-108
Human papillomavirus type-10
Human papillomavirus type-16
Human papillomavirus type-26
Human papillomavirus type-32
Human papillomavirus type-34
Human papillomavirus type-41
Human papillomavirus type-48
Human papillomavirus type-49
Human papillomavirus type-4
Human papillomavirus type-49
Human papillomavirus type-50
Human papillomavirus type-53
Human papillomavirus type-60
Human papillomavirus type-63
Human papillomavirus type 6b
Human papillomavirus type-7
Human papillomavirus type-88
Human papillomavirus type-90
Human papillomavirus type-92
Human papillomavirus type-96
Human papillomavirus type-9
Human parvovirus B19
Human T-lymphotropic virus 1
Human T-lymphotropic virus 4
Hyperthermophilic Archaeal Virus 1
Hyperthermophilic Archaeal Virus 2
Hyphantria cunea nucleopolyhedrovirus
Hyposoter fugitivus ichnovirus segment A1
Hyposoter fugitivus ichnovirus segment A2
Hyposoter fugitivus ichnovirus segment A3
Hyposoter fugitivus ichnovirus segment B10
Hyposoter fugitivus ichnovirus segment B11
Hyposoter fugitivus ichnovirus segment B12
Hyposoter fugitivus ichnovirus segment B13
Hyposoter fugitivus ichnovirus segment B14
Hyposoter fugitivus ichnovirus segment B15
Hyposoter fugitivus ichnovirus segment B16
Hyposoter fugitivus ichnovirus segment B17
Hyposoter fugitivus ichnovirus segment B18
Hyposoter fugitivus ichnovirus segment B19
Hyposoter fugitivus ichnovirus segment B20
Hyposoter fugitivus ichnovirus segment B3
Hyposoter fugitivus ichnovirus segment B4
Hyposoter fugitivus ichnovirus segment B5
Hyposoter fugitivus ichnovirus segment B6
Hyposoter fugitivus ichnovirus segment B7
Hyposoter fugitivus ichnovirus segment B8
Hyposoter fugitivus ichnovirus segment B9
Hyposoter fugitivus ichnovirus segment C10
SIANN: Strain Identification by Alignment to Near Neighbors

Hyposoter fugitivus ichnovirus segment C11
Hyposoter fugitivus ichnovirus segment C12
Hyposoter fugitivus ichnovirus segment C13
Hyposoter fugitivus ichnovirus segment C14
Hyposoter fugitivus ichnovirus segment C15
Hyposoter fugitivus ichnovirus segment C16
Hyposoter fugitivus ichnovirus segment C17
Hyposoter fugitivus ichnovirus segment C18
Hyposoter fugitivus ichnovirus segment C19
Hyposoter fugitivus ichnovirus segment C1
Hyposoter fugitivus ichnovirus segment C20
Hyposoter fugitivus ichnovirus segment C2
Hyposoter fugitivus ichnovirus segment C3
Hyposoter fugitivus ichnovirus segment C4
Hyposoter fugitivus ichnovirus segment C5
Hyposoter fugitivus ichnovirus segment C6
Hyposoter fugitivus ichnovirus segment C7
Hyposoter fugitivus ichnovirus segment C8
Hyposoter fugitivus ichnovirus segment C9
Hyposoter fugitivus ichnovirus segment D10
Hyposoter fugitivus ichnovirus segment D11
Hyposoter fugitivus ichnovirus segment D12
Hyposoter fugitivus ichnovirus segment D1
Hyposoter fugitivus ichnovirus segment D2
Hyposoter fugitivus ichnovirus segment D3

Hyposoter fugitivus ichnovirus segment D4
Hyposoter fugitivus ichnovirus segment D5
Hyposoter fugitivus ichnovirus segment D6
Hyposoter fugitivus ichnovirus segment D7
Hyposoter fugitivus ichnovirus segment D8
Hyposoter fugitivus ichnovirus segment D9
Hyposoter fugitivus ichnovirus segment E1
Hyposoter fugitivus ichnovirus segment E2
Hyposoter fugitivus ichnovirus segment G1

Ictalurid herpesvirus 1 strain Auburn 1
Indian cassava mosaic virus DNA A
Indian cassava mosaic virus DNA B
Infectious hypodermal and hematopoietic necrosis virus
Infectious spleen and kidney necrosis virus
Invertebrate iridescent virus 6
Ipomoea yellow vein virus
Jatropha leaf curl virus DNA A
Jatropha yellow mosaic India virus DNA-A
JC polyomavirus
Junonia coenia densovirus
Kalanchoe top-spotting virus
Kalanchoe phage DNA A
KI polyomavirus Stockholm 60
Klebsiella phage KP15
Klebsiella phage KP32
Klebsiella phage KP34
Klebsiella phage phiKO2
Kluyvera phage Kvp1
Kudzu mosaic virus DNA-A
Kudzu mosaic virus DNA-B
Lactobacillus johnsonii prophage Lj771
Lactobacillus phage A2
Lactobacillus phage KC5a
Lactobacillus phage Lb338-1
Lactobacillus phage Lc-Nu
Lactobacillus phage LL-H
Lactobacillus phage LP65
Lactobacillus phage Lrm1
Lactobacillus phage Lv-1
Lactobacillus phage phiAT3
Lactobacillus phage phiB1e
Lactobacillus phage phiJL-1
Lactobacillus prophage Lj928
Lactobacillus prophage Lj965
Lactobacillus prophage phiadh
Lactococcus phage 1706
Lactococcus phage 4268
Lactococcus phage 712
Lactococcus phage asccphi28
Lactococcus phage bIBB29
Lactococcus phage bil170
Lactococcus phage bil67
Lactococcus phage BK5-T
Lactococcus phage c2
Lactococcus phage jij50
Lactococcus phage KSY1
Lactococcus phage P008
Lactococcus phage P087
Lactococcus phage phiLC3
Lactococcus phage Q54
Lactococcus phage r11
Lactococcus phage sk1
Lactococcus phage TP901-1
Lactococcus phage Tuc2009
Lactococcus phage ul36
Lactococcus prophage bil285
Lactococcus prophage bil286
Lactococcus prophage bil309
Lactococcus prophage bil310
Lactococcus prophage bil311
Lactococcus prophage bil312
Lamium leaf distortion associated virus
Leucania separata nuclear polyhedrosis virus
Leucas zeylanica yellow vein virus satellite DNA beta
<table>
<thead>
<tr>
<th>Virus/Phage Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lindernia anagallis yellow vein virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Lindernia anagallis yellow vein virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Listeria phage A006</td>
<td></td>
</tr>
<tr>
<td>Listeria phage A118</td>
<td></td>
</tr>
<tr>
<td>Listeria phage A500</td>
<td></td>
</tr>
<tr>
<td>Listeria phage A511</td>
<td></td>
</tr>
<tr>
<td>Listeria phage B025</td>
<td></td>
</tr>
<tr>
<td>Listeria phage B054</td>
<td></td>
</tr>
<tr>
<td>Listeria phage P35</td>
<td></td>
</tr>
<tr>
<td>Listeria phage P40</td>
<td></td>
</tr>
<tr>
<td>Listonella phage phiHSIC</td>
<td></td>
</tr>
<tr>
<td>Luffa yellow mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Lucky bamboo bacilliform virus</td>
<td></td>
</tr>
<tr>
<td>Ludwigia leaf distortion betasatellite</td>
<td></td>
</tr>
<tr>
<td>IndiaAmadalavalasaHibiscus2007</td>
<td></td>
</tr>
<tr>
<td>Ludwigia yellow vein virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Ludwigia yellow vein virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Luffa begomovirus associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Luffa puckering and leaf distortion-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Lulli virus</td>
<td></td>
</tr>
<tr>
<td>Lumpy skin disease virus NI-2490</td>
<td></td>
</tr>
<tr>
<td>Lymantria dispar MNPV</td>
<td></td>
</tr>
<tr>
<td>Lymantria xylina MNPV</td>
<td></td>
</tr>
<tr>
<td>Lymphocystis disease virus</td>
<td></td>
</tr>
<tr>
<td>Lymphocystis disease virus isolate China</td>
<td></td>
</tr>
<tr>
<td>Macacine herpesvirus 1</td>
<td></td>
</tr>
<tr>
<td>Macacine herpesvirus 3</td>
<td></td>
</tr>
<tr>
<td>Macacine herpesvirus 4</td>
<td></td>
</tr>
<tr>
<td>Macacine herpesvirus 5</td>
<td></td>
</tr>
<tr>
<td>Macaque simian foamy virus</td>
<td></td>
</tr>
<tr>
<td>Macroptilium golden mosaic virus-</td>
<td></td>
</tr>
<tr>
<td>JamaicaWissadulaAugust</td>
<td></td>
</tr>
<tr>
<td>Town DNA B</td>
<td></td>
</tr>
<tr>
<td>Macroptilium mosaic Puerto Rico virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Macroptilium mosaic Puerto Rico virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Macroptilium yellow mosaic Florida virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Macroptilium yellow mosaic Florida virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Macroptilium yellow mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Macroptilium yellow mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Maize streak virus-ASouth Africa</td>
<td></td>
</tr>
<tr>
<td>Malachra yellow vein mosaic virus-associated satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Mal de Rio Cuarto virus segment 9</td>
<td></td>
</tr>
<tr>
<td>Malvastrum leaf curl Guangdong virus</td>
<td></td>
</tr>
<tr>
<td>Malvastrum leaf curl virus--associated defective DNA beta</td>
<td></td>
</tr>
<tr>
<td>Malvastrum leaf curl virus-G87</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow mosaic virus-associated DNA 1</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow mosaic virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow mosaic virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow vein Baoshan virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow vein-virus</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow vein Yunnan-virus</td>
<td></td>
</tr>
<tr>
<td>Malvastrum yellow vein Yunnan virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td>Mamestra configurata NPV-A</td>
<td></td>
</tr>
<tr>
<td>Mamestra configurata NPV-B</td>
<td></td>
</tr>
<tr>
<td>Mannheimia phage phiMHaA1</td>
<td></td>
</tr>
<tr>
<td>Maruca vitrata MNPV</td>
<td></td>
</tr>
<tr>
<td>Mastomys coucha papillomavirus 2</td>
<td></td>
</tr>
<tr>
<td>Mastomys natalensis papillomavirus</td>
<td></td>
</tr>
<tr>
<td>Melanoplus sanguinipes entomopoxvirus</td>
<td></td>
</tr>
<tr>
<td>Meleagrid herpesvirus 1</td>
<td></td>
</tr>
<tr>
<td>Melon chlorotic leaf curl virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Melon chlorotic mosaic virus-associated alphasatellite</td>
<td></td>
</tr>
<tr>
<td>Melon chlorotic mosaic virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Melon chlorotic mosaic virus DNA-B</td>
<td></td>
</tr>
<tr>
<td>Merkel cell polyomavirus</td>
<td></td>
</tr>
<tr>
<td>Merremia mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Merremia mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Mesta yellow vein mosaic Bahraich virus-IndiaBahraich2007 DNA A</td>
<td></td>
</tr>
<tr>
<td>Mesta yellow vein mosaic virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td>Mesta yellow vein mosaic virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Methanobacterium phage psm2</td>
<td></td>
</tr>
<tr>
<td>Methanothermobacter prophage psm100</td>
<td></td>
</tr>
<tr>
<td>Microbacterium phage Min1</td>
<td></td>
</tr>
<tr>
<td>Microcystis phage Ma-LMM01</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment A</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment B</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment C</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment D</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment E</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment F</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment G</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment H</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment I</td>
<td></td>
</tr>
<tr>
<td>Microplitis bracovirus segment J</td>
<td></td>
</tr>
</tbody>
</table>

This list is a sample of viruses identified by the SIANN method. The method aligns strains to near neighbors to identify new virus species and strains.
SIANN: Strain Identification by Alignment to Near Neighbors

- Microplitis demolitor bracovirus segment L
- Microplitis demolitor bracovirus segment M
- Microplitis demolitor bracovirus segment N
- Microplitis demolitor bracovirus segment O
- Milk vetch dwarf virus segment 10
- Milk vetch dwarf virus segment 11
- Milk vetch dwarf virus segment 1
- Milk vetch dwarf virus segment 2
- Milk vetch dwarf virus segment 3
- Milk vetch dwarf virus segment 4
- Milk vetch dwarf virus segment 5
- Milk vetch dwarf virus segment 6
- Milk vetch dwarf virus segment 7
- Milk vetch dwarf virus segment 8
- Milk vetch dwarf virus segment 9
- Mimosa yellow leaf curl virus-associated DNA 1
- Mimosa yellow leaf curl virus DNA-A
- Mimosa yellow leaf curl virus satellite DNA beta
- Minute virus of mice
- Mirabilis mosaic virus
- Miscanthus streak virus-91
- Molluscum contagiosum virus subtype 1
- Monkeypox virus Zaire-96-I-16
- Morganella phage MmP1
- Mouse mammary tumor virus
- Mouse parvovirus 1
- Mouse parvovirus 2
- Mouse parvovirus 3
- Mouse parvovirus 4
- Mouse parvovirus 5
- Mulard duck circovirus
- Mungbean yellow mosaic India virus DNA A
- Mungbean yellow mosaic India virus DNA B
- Mungbean yellow mosaic virus DNA A
- Mungbean yellow mosaic virus DNA B
- Murid herpesvirus 1
- Murid herpesvirus 2
- Murid herpesvirus 4
- Murine adenovirus 3
- Murine adenovirus A
- Murine pneumotropic virus
- Murine polyomavirus
- Murine type C retrovirus
- Musca domestica salivary gland hypertrophy virus
- Muscovy duck circovirus
- Muscovy duck parovirus
- Mus musculus papillomavirus type 1
- Mycobacterium phage 244
- Mycobacterium phage Adjutor
- Mycobacterium phage Angel
- Mycobacterium phage angelica
- Mycobacterium phage Ardmore
- Mycobacterium phage Barnyard
- Mycobacterium phage Bethlehem
- Mycobacterium phage Boomer
- Mycobacterium phage BPs
- Mycobacterium phage Brujita
- Mycobacterium phage Butterscotch
- Mycobacterium phage Bxb1
- Mycobacterium phage Bxz1
- Mycobacterium phage Bxz2
- Mycobacterium phage Cali
- Mycobacterium phage Catera
- Mycobacterium phage Chah
- Mycobacterium phage Che12
- Mycobacterium phage Che8
- Mycobacterium phage Che9c
- Mycobacterium phage Che9d
- Mycobacterium phage Cjw1
- Mycobacterium phage Cooper
- Mycobacterium phage Corndog
- Mycobacterium phage CrimD
- Mycobacterium phage D29
- Mycobacterium phage DD5
- Mycobacterium phage ET08
- Mycobacterium phage Fruitloop
- Mycobacterium phage Giles
- Mycobacterium phage Gumball
- Mycobacterium phage Halo
- Mycobacterium phage Jasper
- Mycobacterium phage KBG
- Mycobacterium phage L5
- Mycobacterium phage Konstantine
- Mycobacterium phage Kostya
- Mycobacterium phage Lij
- Mycobacterium phage Lockley
- Mycobacterium phage Myrna
- Mycobacterium phage Nigel
- Mycobacterium phage Omega
- Mycobacterium phage Orion
- Mycobacterium phage Pacc40
- Mycobacterium phage Peaches
- Mycobacterium phage PG1
- Mycobacterium phage Phaedrus
- Mycobacterium phage Phlyer
- Mycobacterium phage Pipefish
- Mycobacterium phage Plot
- Mycobacterium phage PMC
SIANN: Strain Identification by Alignment to Near Neighbors

Mycobacterium phage Porky
Mycobacterium phage Predator
Mycobacterium phage Pukovnik
Mycobacterium phage Qyrzula
Mycobacterium phage Ramsey
Mycobacterium phage Rizal
Mycobacterium phage Rosebush
Mycobacterium phage ScottMcG
Mycobacterium phage Solon
Mycobacterium phage Spud
Mycobacterium phage TM4
Mycobacterium phage Troll4
Mycobacterium phage Tweety
Mycobacterium phage U2
Mycobacterium phage Wildcat
Mycoplasma phage MAV1
Mycoplasma phage P1
Mycoplasma phage phiMFV1
Myotis polymavirus VM-2008
Myxococcus phage Mx8
Myxoma virus
Myzus persicae densovirus
Nanovirus-like particle
Natritalpha phage PhiCh1
Neodiprion abietis NPV
Neodiprion lecontei NPV
Neodiprion sertifer NPV
Oat dwarf virus
OklLCV satDNA 10
Okra leaf curl disease associated DNA 1
Okra leaf curl Mali virus satellite DNA beta
Okra leaf curl virus-Cameroon
Okra mottle virus-Brazilokra DNA A
Okra mottle virus-Brazilokra DNA B
Okra yellow crinkle virus segment A
Okra yellow mosaic Mexico virus DNA A
Okra yellow mosaic Mexico virus DNA B
Okra yellow vein disease associated sequence virion
Okra yellow vein mosaic virus
Old World harvest mouse papillomavirus
Orangutan polymavirus
Orf virus
Orgyia leucostigma NPV
Orgyia pseudotsugata MNPV
Oryctes rhinoceros virus
Ostreid herpesvirus 1
Ostreococcus tauri virus 1
Ostreococcus virus OsV5
Ovine adenovirus A
Ovine adenovirus D
Ovine herpesvirus 2
Ovine papillomavirus-1
Panicum streak virus-Karino
Panine herpesvirus 2 strain Heberling
Papaya leaf curl China virus-G8
Papaya leaf curl China virus satellite DNA beta
Papaya leaf curl Guandong virus-GD2 DNA A
Papaya leaf curl virus-associated DNA beta
Papaya leaf curl-virus
Papiine herpesvirus 2
Paramacium bursaria
Chlorella virus 1
Paramacium bursaria
Chlorella virus AR158
Paramacium bursaria
Chlorella virus FR483
Paramacium bursaria
Chlorella virus NY2A
Parovirus H1
Passionfruit severe leaf distortion virus DNA-A
Passionfruit severe leaf distortion virus DNA-B
Pasteurella phage FI08
Peanut chlorotic streak virus
Pedilanthus leaf curl virus-Pedilanthus
PakistanMultan2004
Pelargonium vein banding virus
Penaeus merguiensis densovirus
Penaeus monodon hepatopancreatic parovirus
Pepper curly top virus
Pepper golden mosaic virus DNA A
Pepper golden mosaic virus DNA B
Pepper huasteco yellow vein virus DNA A
Pepper huasteco yellow vein virus DNA B
Pepper leaf curl Bangladesh virus segment A component
Pepper leaf curl virus DNA-A
Pepper leaf curl virus satellite DNA beta
Pepper leaf curl Yunnan virus satellite DNA beta
Pepper leaf curl Yunnan virus-YN323
Pepper yellow dwarf virus-New Mexico
Pepper yellow leaf curl Indonesia virus DNA-A
Pepper yellow leaf curl Indonesia virus DNA-B
Pepper yellow vein Mali virus
Periplaneta fuliginosa densovirus
Petunia vein clearing virus
Phage cd11
Phage Gifsy-1
Phage Gifsy-2
Phage phiJL001
Phocoena spinipinnis papillomavirus
Phormidium phage Pf-WMP3
Phormidium phage Pf-WMP4
Phthorimaea operculella granulovirus
Pieris rapae granulovirus
Planaria asexual strain-specific virus-like element type 1 large DNA segment
Planaria asexual strain-specific virus-like element type 1 small DNA segment
SIANN: Strain Identification by Alignment to Near Neighbors

Planococcus citri densovirus
Plutella xylostella granulovirus
Plutella xylostella multiple nucleopolyhedrovirus
Polyomavirus HPyV6
Polyomavirus HPyV7
Porcine adenovirus C
Porcine circovirus 1
Porcine circovirus 2
Porcine endogenous retrovirus E
Porcine parvovirus
Potato apical leaf curl disease-associated satellite DNA beta
Potato yellow mosaic Panama virus DNA A
Potato yellow mosaic Panama virus DNA B
Potato yellow mosaic Trinidad virus DNA A
Potato yellow mosaic Trinidad virus DNA B
Potato yellow mosaic virus DNA A
Potato yellow mosaic virus DNA B
Prochlorococcus phage P-SSM4
Propionibacterium phage B5
Propionibacterium phage PA6
Pseudaletia unipuncta granulovirus
Pseudoalteromonas phage PM2
Pseudocowpox virus
Pseudomonas phage 119X
Pseudomonas phage 14-1
Pseudomonas phage 201phi2-1
Pseudomonas phage 73
Pseudomonas phage B3
Pseudomonas phage D3112
Pseudomonas phage-D3
Pseudomonas phage DM53
Pseudomonas phage EL
Pseudomonas phage F10
Pseudomonas phage F116
Pseudomonas phage F8
Pseudomonas phage gh-1
Pseudomonas phage LBL3
Pseudomonas phage LIT1
Pseudomonas phage LKD16
Pseudomonas phage LMA2
Pseudomonas phage LUZ19
Pseudomonas phage LUZ24
Pseudomonas phage LUZ7
Pseudomonas phage M6
Pseudomonas phage MP22
Pseudomonas phage MP29
Pseudomonas phage MP38
Pseudomonas phage PA11
Pseudomonas phage PAJU2
Pseudomonas phage PaP2
Pseudomonas phage phage PaP3
Pseudomonas phage PB1
Pseudomonas phage Pf1
Pseudomonas phage Pf3
Pseudomonas phage phi-2
Pseudomonas phage phiCTX
Pseudomonas phage phiK77
Pseudomonas phage phiIMV
Pseudomonas phage phiKZ
Pseudomonas phage PT2
Pseudomonas phage PT5
Pseudomonas phage SN
Pseudomonas phage YuA
Psittacid herpesvirus 1
Psittacus erithacus timneh papillomavirus
Pumpkin yellow mosaic Malaysia virus DNA A
Pyrobaculum spherical virus
Pyrococcus abyssi virus 1
Rabbit fibroma virus
Rabbit oral papillomavirus
Rachiplusia ou MNPV
Radish leaf curl virus satellite DNA beta
Radish leaf curl virus segment A
Ralstonia phage p12J
Ralstonia phage phiRSA1
Ralstonia phage RS81
Ralstonia phage RSL1
Ralstonia phage RSM1
Ralstonia phage RSM3
Ralstonia phage RSS1
Ramie mosaic virus DNA-A
Ramie mosaic virus DNA-B
Ranid herpesvirus 1 strain McKinnell
Ranid herpesvirus 2 strain ATCC VR-568
Rauscher murine leukemia virus
Raven circovirus
RD114 retrovirus
Reticuloendotheliosis virus
Rhesus monkey papillomavirus
Rhizobium phase 16-3
Rhodothermus phase RM378
Rhynchosia golden mosaic virus DNA A
Rhynchosia golden mosaic virus DNA B
Rhynchosia golden mosaic virus Yucatan virus DNA A
Rhynchosia golden mosaic virus Yucatan virus DNA B
Rice tungro bacilliform virus
Roseobacter phase SIO1
Roseophage DSS3P2
Roseophage EE36P1
Rousettus aegyptiacus papillomavirus type 1
Rudbeckia flower distortion virus
Saccharum streak virus
Saimiriine herpesvirus 2
Salmonella enterica bacteriophage SE1
Salmonella phage c341
Salmonella phage E1
Salmonella phage epsilon34
Salmonella phage Fels-1
Salmonella phage phiSG-JL2
Salmonella phage SETP3
Salmonella phage St64B
Sclerotinia sclerotiorum hypovirulence associated DNA virus 1
Sea turtle tornovirus 1
Senecio yellow mosaic virus
Sheeppox virus 17077-99
Sheldgoose hepatitis B virus
Shigella phage phiSboM-AG3
Shrimp white spot syndrome virus
SIANN: Strain Identification by Alignment to Near Neighbors

Sida golden mosaic Costa Rica virus DNA A
Sida golden mosaic Costa Rica virus DNA B
Sida golden mosaic Florida virus-Malvastrum DNA-A
Sida golden mosaic Florida virus-Malvastrum DNA-B
Sida golden mosaic Honduras virus DNA A
Sida golden mosaic Honduras virus DNA B
Sida golden mosaic virus DNA A
Sida golden mosaic virus DNA-B
Sida golden mottle virus DNA-A
Sida golden mottle virus DNA-B
Sida leaf curl virus-associated DNA 1
Sida leaf curl virus-associated DNA beta
Sida leaf curl virus satellite DNA beta
Sida micrantha mosaic virus segment A
Sida micrantha mosaic virus segment B
Sida mosaic Sinaloa virus DNA A
Sida mosaic Sinaloa virus DNA B
Sida mottle virus
Sida yellow mosaic virus-China-associated DNA beta DNA beta
Sida yellow mosaic-virus
Sida yellow mosaic Yucatan virus DNA A
Sida yellow mosaic Yucatan virus DNA B
Sida yellow vein disease associated DNA 1
Sida yellow vein Madurai virus
Sida yellow vein Vietnam virus-associated DNA 1
Sida yellow vein Vietnam virus DNA-A
Sida yellow vein Vietnam virus satellite DNA beta
Sida yellow vein virus DNA A
Sida yellow vein virus DNA B
Sida yellow vein virus satellite DNA beta
Siegesbeckia yellow vein Guanxi virus
Siegesbeckia yellow vein virus-GD13-associated DNA beta
Siegesbeckia yellow vein virus GD13
Simian adenovirus 1
Simian adenovirus 3
Simian foamy virus 3
Simian immunodeficiency virus SIV-mnd 2
Simian retrovirus 4
Simian T-cell lymphotropic virus 6
Simian T-lymphotropic virus 1
Simian T-lymphotropic virus 3
Simian virus 40
Singapore grouper iridovirus
Sinorhizobium phage PBC5
Small anellovirus 1
Small anellovirus 2
Snake parovirus 1
Snow goose hepatitis B virus
Sodalis phage phiSG1
Sodalis phage SO-1
South African cassava mosaic virus DNA A
South African cassava mosaic virus DNA B
Soybean chlorotic blotch virus DNA A
Soybean chlorotic blotch virus DNA B
Soybean chlorotic mottle virus
Soybean crinkle leaf virus
Soybean mild mottle virus
Spilanthes yellow vein virus DNA-A
Spinach curly top virus
Spiroplasma kunkeli virus SkV1 CR2-3x
Spiroplasma phage 1-C74
Spiroplasma phage 1-R8A2B
Spiroplasma phage 4
Spiroplasma phage SVTS2
Spodoptera exigua MNPV
Spodoptera frugiperda ascovirus 1a
Spodoptera frugiperda MNPV virus
Spodoptera litura granulovirus
Spodoptera litura NPV
Spodoptera litura nucleopolyhedrovirus II
Sputnik virophage
Squash leaf curl China virus-B DNA-A
Squash leaf curl China virus-B DNA B
Squash leaf curl Philippines virus segment A
Squash leaf curl Philippines virus segment B
Squash leaf curl virus A component DNA
Squash leaf curl virus B component DNA
Squash leaf curl Yunnan virus
Squash mild leaf curl virus-Imperial Valley DNA A
Squash mild leaf curl virus-Imperial Valley DNA B
Squash yellow mild mottle virus DNA B
Squirrel monkey polyomavirus
Sri Lankan cassava mosaic virus DNA A
Sri Lankan cassava mosaic virus DNA B
Stachytyrpheta leaf curl virus
Staphylococcus phage 11
Staphylococcus phage 187
Staphylococcus phage 2638A
Staphylococcus phage 29
Staphylococcus phage 37
Staphylococcus phage 3A
Staphylococcus phage 42e
Staphylococcus phage 44AHJD
Staphylococcus phage 47
Staphylococcus phage 52A
Staphylococcus phage 53
Staphylococcus phage 55
Staphylococcus phage 66
Staphylococcus phage 69
| Staphylococcus phage 71 | Staphylococcus phage 77 | Staphylococcus phage 80alpha | Staphylococcus phage 85 | Staphylococcus phage 88 | Staphylococcus phage 92 | Staphylococcus phage 96 | Staphylococcus phage CNPH82 | Staphylococcus phage EW | Staphylococcus phage G1 | Staphylococcus phage K | Staphylococcus phage P954 | Staphylococcus phage PH15 | Staphylococcus phage phi2958PVL | Staphylococcus phage phiETA2 | Staphylococcus phage phiETA3 | Staphylococcus phage phiETA | Staphylococcus phage phiMR11 | Staphylococcus phage phiMR25 | Staphylococcus phage phiNM1 | Staphylococcus phage phiNM3 | Staphylococcus phage phiP68 | Staphylococcus phage phiPVL108 | Staphylococcus phage phiPVL-CN125 | Staphylococcus phage phiSauS-IPLA35 | Staphylococcus phage phiSauS-IPLA88 | Staphylococcus phage phiSLT | Staphylococcus prophage phi12 | Staphylococcus prophage phi13 | Staphylococcus prophage phiN315 | Staphylococcus prophage phiPV83 | Staphylococcus prophage PVL | Staphylococcus prophage tp310-1 | Staphylococcus prophage tp310-2 | Staphylococcus prophage tp310-3 | Starling circovirus | Stenotrophomonas phage phiSMA9 | Stenotrophomonas phage S1 | Strawberry vein banding virus | Streptococcus phage 2972 | Streptococcus phage 5093 | Streptococcus phage 7201 | Streptococcus phage 858 | Streptococcus phage Abc2 | Streptococcus phage ALQ13.2 | Streptococcus phage C1 | Streptococcus phage Cp-1 | Streptococcus phage DT1 | Streptococcus phage M102 | Streptococcus phage O1205 | Streptococcus phage P9 | Streptococcus phage PH10 | Streptococcus phage PH15 | Streptococcus phage phi3396 | Streptococcus phage phi11 | Streptococcus phage phi19 | Streptococcus phage phi21 | Streptococcus phage SM1 | Streptococcus prophage SMP | Streptococcus prophage 315.1 | Streptococcus prophage 315.2 | Streptococcus prophage 315.5 | Streptococcus prophage 315.6 | Streptococcus prophage EJ-1 | Streptococcus prophage MM1 | Streptococcus pyogenes phage 315.3 | Streptomyces phage mu16 | Streptomyces phage phiBT1 | Streptomyces phage phiIC31 | Streptomyces phage phiSASD1 | Streptomyces phage VWB | Stx1 converting phase | Stx2-converting phase 1717 | Stx2-converting phase 86 | Stx2 converting phase I | Stx2 converting phase II | Subterranean clover stunt virus DNA 1 | Subterranean clover stunt virus DNA 2 | Subterranean clover stunt virus DNA 3 | Subterranean clover stunt virus DNA 4 | Subterranean clover stunt virus DNA 5 | Subterranean clover stunt virus DNA 6 | Subterranean clover stunt virus DNA 7 | Subterranean clover stunt virus DNA 8 | Sugarcane bacilliform IM virus | Sugarcane bacilliform Mor virus | Sugarcane bacilliform virus | Sugarcane streak Egypt virus-Giza | Sugarcane streak Reunion virus | Sugarcane streak virus-Natal | Suid herpesvirus 1 | Sulfolobus islandicus filamentous virus | Sulfolobus islandicus rod-shaped virus 1 | Sulfolobus islandicus rod-shaped virus 2 | Sulfolobus spindle-shaped virus 4 | Sulfolobus spindle-shaped virus 5 |
SIANN: Strain Identification by Alignment to Near Neighbors

Sulfolobus spindle-shaped virus 6
Sulfolobus spindle-shaped virus 7
Sulfolobus turreted icosahedral virus 2
Sulfolobus turreted icosahedral-virus 1
Sulfolobus virus 2
Sulfolobus virus Kamchatka 1
Sulfolobus virus Ragged Hills
Sulfolobus virus STSV1
Sunn hemp leaf distortion virus DNA-A
Sus scrofa papillomavirus type 1
Sweetpotato badnavirus B
Sweet potato leaf curl Bengal virus-India West Bengal2008 segment A
Sweet potato leaf curl Canary virus
Sweet potato leaf curl Georgia virus
Sweet potato leaf curl Lanzarote virus
Sweet potato leaf curl Spain virus
Sweet potato leaf curl virus
Swinepox virus
Synechococcus phage P60
Synechococcus phage S-PM2
Synechococcus phage S-RSM4
Synechococcus phage syn9
Tanapox virus
Taro bacilliform virus
Taterapox virus
Thalassomonas phage BA3
Thermoproteus tenax spherical virus 1
Thermus phage IN93
Thermus phage P23-45
Thermus phage P23-77
Thermus phage P74-26
Thermus phage phiYS40
Tobacco curly shoot virus associated DNA 1
Tobacco curly shoot virus-associated DNA beta
Tobacco curly shoot-virus
Tobacco leaf curl disease associated sequence virion
Tobacco leaf curl Japan virus
Tobacco leaf curl Kochi virus
Tobacco leaf curl Thailand virus
Tobacco leaf curl virus-associated DNA beta
Tobacco leaf curl Yunnan virus associated DNA 1
Tobacco leaf curl Yunnan virus satellite DNA beta
Tobacco leaf curl Yunnan virus-Y136
Tobacco leaf curl Zimbabwe virus
Tobacco vein clearing virus
Tobacco yellow dwarf virus
Tomato begomovirus satellite DNA beta
Tomato chino La Paz virus segment A
Tomato chlorotic mottle virus DNA A
Tomato chlorotic mottle virus DNA B
Tomato common mosaic virus DNA-A
Tomato common mosaic virus DNA-B
Tomato curly stunt virus
Tomato golden mosaic virus DNA A
Tomato golden mosaic virus DNA B
Tomato golden mottle virus DNA A
Tomato golden mottle virus DNA B
Tomato leaf curl Arusha virus DNA-A
Tomato leaf curl Bangalore virus-Ban5 satellite DNA beta
Tomato leaf curl Bangalore-virus
Tomato leaf curl Bangladesh virus
Tomato leaf curl Cameroon virus-
CameroonBueaOkra2008
Tomato leaf curl Cebu virus
Tomato leaf curl China virus-G32
Tomato leaf curl Cotabato virus DNA-A
Tomato leaf curl Ghana virus segment A
Tomato leaf curl Guangdong virus DNA-A
Tomato leaf curl Guangxi virus
Tomato leaf curl Gujarat virus-Varanasi segment A
Tomato leaf curl Gujarat virus-Varanasi segment B
Tomato leaf curl Hainan virus
Tomato leaf curl Hsinchu virus-TaiwanHsinchu2005 DNA A
Tomato leaf curl Iran virus
Tomato leaf curl Java virus-Ageratum satellite DNA
Tomato leaf curl Java-virus
Tomato leaf curl Joydebpur beta virus
Tomato leaf curl Joydebpur virus DNA-A
Tomato leaf curl Karnataka virus-associated DNA beta DNA-A
Tomato leaf curl Karnataka-virus
Tomato leaf curl Kerala virus
Tomato leaf curl Kumasi virus segment A
Tomato leaf curl Laos virus
Tomato leaf curl Malaysia virus
Tomato leaf curl Mali virus
Tomato leaf curl Mayotte virus
Tomato leaf curl Mindanao virus DNA-A
Tomato leaf curl New Delhi virus-associated DNA beta
Tomato leaf curl New Delhi virus DNA A
Tomato leaf curl New Delhi virus DNA B
Tomato leaf curl Nigeria virus-Nigeria2006
Tomato leaf curl Pakistan virus associated DNA 1
SIANN: Strain Identification by Alignment to Near Neighbors

- Tomato leaf curl Pakistan virus
  - segment A
- Tomato leaf curl Palampur virus
- Tomato leaf curl Patna virus
- DNA-A
- Tomato leaf curl Philippines virus
- Tomato leaf curl Philippine virus satellite DNA beta
- Tomato leaf curl Pune virus
- Tomato leaf curl Seychelles virus
- Tomato leaf curl Sinaloa virus DNA-A
- Tomato leaf curl Sinaloa virus DNA-B
- Tomato leaf curl Sri Lanka virus
- Tomato leaf curl Sudan virus-Gezira
- Tomato leaf curl Sulawesi virus DNA-A
- Tomato leaf curl Taiwan virus
- Tomato leaf curl Togo virus-
  - Togo2006
- Tomato leaf curl Vietnam virus DNA-A
- Tomato leaf curl virus-associated DNA beta
- Tomato leaf curl-virus
- Tomato leaf curl virus-Pune-associated DNA beta DNA-A
- Tomato mild mosaic virus DNA-A
- Tomato mild mosaic virus DNA-B
- Tomato mild yellow leaf curl
  - Aragua virus DNA-A
- Tomato mild yellow leaf curl
  - Aragua virus DNA-B
- Tomato mosaic Havana virus DNA-A
- Tomato mosaic Havana virus DNA-B
- Tomato mosaic leaf curl virus DNA-A
- Tomato mosaic leaf curl virus DNA-B
- Tomato mottle Taino virus DNA-A
- Tomato mottle Taino virus DNA-B
- Tomato mottle virus DNA-A
- Tomato mottle virus DNA-B
- Tomato pseudo-curl leaf curl virus
- Tomato rugose mosaic virus DNA-A
- Tomato rugose mosaic virus DNA-B
- Tomato severe leaf curl virus
- Tomato severe rugose virus DNA-A
- Tomato severe rugose virus DNA-B
- Tomato yellow dwarf disease associated satellite DNA beta-Kochi virus
- Tomato yellow leaf curl China virus associated DNA 1
- Tomato yellow leaf curl China-virus
- Tomato yellow leaf curl Guangdong virus DNA-A
- Tomato yellow leaf curl Indonesia virus-Lembang
- Tomato yellow leaf curl Kanchanaburi virus DNA-A
- Tomato yellow leaf curl Kanchanaburi virus DNA-B
- Tomato yellow leaf curl Mali virus-associated DNA beta
- Tomato yellow leaf curl Thailand virus associated DNA 1
- Tomato yellow leaf curl Thailand virus DNA-A
- Tomato yellow leaf curl Thailand virus DNA-B
- Tomato yellow leaf curl Vietnam virus DNA-A
- Tomato yellow leaf curl Vietnam virus satellite DNA beta
- Tomato yellow leaf curl virus-associated DNA beta
- Tomato yellow margin leaf curl virus DNA-A
- Tomato yellow margin leaf curl virus DNA-B
- Tomato yellow spot virus DNA-A
- Tomato yellow spot virus DNA-B
- Tomato yellow vein streak virus DNA-A
- Tomato yellow vein streak virus DNA-B
- Torque teno canis virus
- Torque teno dourocouli virus
- Torque teno felis virus
- Torque teno midi virus 1
- Torque teno midi virus 2
- Torque teno mini virus 1
- Torque teno mini virus 2
- Torque teno mini virus 3
- Torque teno mini virus 4
- Torque teno mini virus 5
- Torque teno mini virus 6
- Torque teno mini virus 7
- Torque teno mini virus 8
- Torque teno mini virus 9
- Torque teno sus virus 1
- Torque teno tamarin virus
- Torque teno virus 10
- Torque teno virus 12
- Torque teno virus 14
- Torque teno virus 15
- Torque teno virus 16
- Torque teno virus 19
- Torque teno virus-1
- Torque teno virus 25
- Torque teno virus 26
- Torque teno virus 27
- Torque teno virus 28
- Torque teno virus 3
- Torque teno virus 4
- Torque teno virus 6
- Torque teno virus 7
- Torque teno virus 8
- Trichodysplasia spinulosa-associated polyomavirus
- Trichoplusia ni ascovirus 2c
- Trichoplusia ni SNPV
- Tuapiaid herpesvirus 1
- Turkey adenovirus A
- Turnip curly top virus
- TYLCCNV-Y322 satellite DNA beta
- Urochloa streak virus
- Vaccinia virus
- Variola virus
SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Virus/Phage Name</th>
<th>Virus/Phage Name</th>
<th>Virus/Phage Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velvet bean severe mosaic virus DNA A</td>
<td>Vibrio phage VP2</td>
<td>Xanthomonas phage phiL7</td>
</tr>
<tr>
<td>Velvet bean severe mosaic virus DNA B</td>
<td>Vibrio phage VP4</td>
<td>Xanthomonas phage Xop411</td>
</tr>
<tr>
<td>Vernonia yellow vein betasatellite DNA A</td>
<td>Vibrio phage VP5</td>
<td>Xanthomonas phage Xp10</td>
</tr>
<tr>
<td>Vernonia yellow vein virus DNA-A</td>
<td>Vibrio phage VP882</td>
<td>Xanthomonas phage Xp15</td>
</tr>
<tr>
<td>Vibrio phage fs1</td>
<td>Vibrio phage VP93</td>
<td>Xenopus laevis endogenous retravirus Xen1</td>
</tr>
<tr>
<td>Vibrio phage fs2</td>
<td>Vibrio phage VpV262</td>
<td>Xestia c-nigrum granulovirus</td>
</tr>
<tr>
<td>Vibrio phage K139</td>
<td>Vibrio phage VSK</td>
<td>Xylella phage Xfas53</td>
</tr>
<tr>
<td>Vibrio phage kappa</td>
<td>Watermelon chlorotic stunt virus DNA A</td>
<td>Yaba-like disease virus</td>
</tr>
<tr>
<td>Vibrio phage KSF-1phi</td>
<td>Watermelon chlorotic stunt virus DNA B</td>
<td>Yaba monkey tumor virus</td>
</tr>
<tr>
<td>Vibrio phage N4</td>
<td>Wheat dwarf virus</td>
<td>Yersinia pestis phage phiA122</td>
</tr>
<tr>
<td>Vibrio phage VEJphi</td>
<td>Wissadula golden mosaic St</td>
<td>Yersinia phage Berlin</td>
</tr>
<tr>
<td>Vibrio phage VF12</td>
<td>Thomas Virus DNA A</td>
<td>Yersinia phage L-413C</td>
</tr>
<tr>
<td>Vibrio phage VF03K6</td>
<td>Wissadula golden mosaic St</td>
<td>Yersinia phage phiYeO3-12</td>
</tr>
<tr>
<td>Vibrio phage VF04K68</td>
<td>Thomas Virus DNA B</td>
<td>Yersinia phage PY54</td>
</tr>
<tr>
<td>Vibrio phage VGJphi</td>
<td>Woodchuck hepatitis virus</td>
<td>Zinnia leaf curl disease</td>
</tr>
<tr>
<td>Vibrio phage VHML</td>
<td>WU Polyomavirus</td>
<td>associated sequence virion</td>
</tr>
<tr>
<td></td>
<td>Xanthomonas phage Cf1c</td>
<td>Zinnia leaf curl virus-associated DNA beta</td>
</tr>
<tr>
<td></td>
<td>Xanthomonas phage OP1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xanthomonas phage OP2</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix 3: Bacterial Database

<table>
<thead>
<tr>
<th>Bacterial Name</th>
<th>Accession</th>
<th>Strain Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acaryochloris marina MBIC1 1017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetobacter pasteurianus IFO</td>
<td>3283 01</td>
<td></td>
</tr>
<tr>
<td>Acetohalobium arabaticum DSM 5501</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acholeplasma laidlawii PG 8A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achromobacter xylosoxidans A8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidaminococcus fermentans DSM 20731</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidaminococcus intestini RyC MR95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidianus hospitalis W1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidibacter saccharovorans 345 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidimicrobium ferrooxidans DSM 10331</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidiphilium cryptum JF 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidiphilium multivorum AIU301</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidithiobacillus caldus SM 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidithiobacillus ferrivorans SS3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidithiobacillus ferrooxidans ATCC 23270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidobacterium capsulatum ATCC 51196</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidothermus cellulolyticus 11B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidovorax avenae subsp avenae ATCC 19860</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidovorax citrulli AAC00 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidovorax ebreus TPSY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acidovorax sp JS42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aciduliprofundum boonei T469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acinetobacter baumannii 1656 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acinetobacter calcoaceticus PHEA 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acinetobacter oleivorans DR1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acinetobacter sp ADP1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinobacillus pleuropneumoniae serovar 5b sp L20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinobacillus succinogenes 130Z</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinoplanes sp SE50 110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinosynnema mirum DSM 43827</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerococcus urinae ACS 120 V Col10a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aeromonas hydrophila subsp hydrophila ATCC 7966</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aeromonas salmonicida subsp salmonicida A449</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aeropyrum pernix K1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregatibacter actinomycetemcomitans D115 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregatibacter aphrophilus NJ8700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agrobacterium fabrum subsp C58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agrobacterium radiobacter K84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agrobacterium sp H13 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agrobacterium viliis S4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akkermansia muciniphila ATCC BAA 835</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcanivorax borkumensis SK2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alicyclobacillus denitrificans BC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alicyclobacillus acidocaldarius subsp acidocaldarius DSM 446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aliviibrio salmonicida LF1238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkalilimnicola ehrlichii MLHE 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkaliphilus metalliredigens QYMF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkaliphilus orelandii OHILAs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allochromatium vinosum DSM 180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alteromonas macleodii str Deep ecotype</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alteromonas sp SN2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aminobacterium colombiense DSM 12261</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammonifex degensii KC4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amycolatopsis mediterranei U32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amycolicoccus subflavus DQ53 9A1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaebaena variabilis ATCC 29413</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerococcus prevotii DSM 20548</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerolinea thermophila UNI 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaeromyxobacter dehalogenans 2CP 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaeromyxobacter sp Fw109 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaplasma centrale str Israel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaplasma marginale str Florida</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaplasma phagocytophilum HZ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaoybacillus flavithermus WK1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aquifex aeolicus VF5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcanobacterium haemolyticum DSM 20595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaeoglobus fulgidus DSM 4304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaeoglobus profundus DSM 5631</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaeoglobus veneficus SNP6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcobacter butzleri RM4018</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcobacter nitrofigilis DSM 7299</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcobacter sp L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aromatoleum aromaticum EbN1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter arlaitensis Re117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter aurescens TC1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter chlorophenolicus A6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter phenanthrenivorans Sphe3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter sp FB24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aster yellows witches broom phytoplasma AYW8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asticcacaulis excentricus C8 48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atopobium parvulum DSM 20469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Azotobacter vinelandii DJ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus anthracis str Ames</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus amyloliquefaciens DSM 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus acidocaldarius DSM 446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus acidocaldarius subsp acidocaldarius DSM 446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organism</td>
<td>Strain Identifier</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Bacillus atrophaeus 1942</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus cellulosilyticus DSM 2522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus cereus 03BB102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus clausii KSM K16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus coagulans 2 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus cytotoxicus NVH 391 98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus halodurans C 125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus licheniformis DSM 13 ATCC 14580</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus megaterium DSM 319</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus mycoides DSM 2048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus pseudofirmus OF4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus pumilus SAFR 032</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus selenitireducens MLS10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus subtilis BSn5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus thuringiensis serovar berliner ATCC 10792</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillus weihenstephanensis KBA64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteriovorax marinus SJ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides fragilis YCH46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides helcogenes P 36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides salanitronis DSM 18170</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides thetaotaomicron VPI 5482</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroides vulgatus ATCC 8482</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella bacilliformis KC583</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella clarridgeiae 73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella grahamii as4aup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella henselae str Houston 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella quintana strain Toulouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bartonella tribocorum CIP 105476</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baumannia cicadellinicola str Hc Homalodisc a coagulata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bdellovibrio bacteriovorus HD100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beijerinckia indica subspecies indica ATCC 9039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beutenbergia cavernae DSM 12333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium adolescentis ATCC 15703</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium animalis subsp animalis ATCC 25527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium bifidum S17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium breve ACS 071 V Sch8b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium denditium Bd1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium longum DJO10A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blattabacterium sp Blattella germanica str Bge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella avium 197N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella bronchiseptica RB50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella parapertussis 12822</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella pertussis CS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bordetella petii DSM 12804</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia afzeli PKo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia bissetti DN127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia burgdorferi B31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia duttonii Ly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia garinii PBI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia hermsii DAH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia recurrentis A1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borrelia turicata 91E135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachybacterium faecium DSM 4810</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachyspira hyodysenteriae WA1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachyspira intermedia PWS A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachyspira muralochii DSM 12563</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachyspira pilosicoli 95 1000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bradyrhizobium japonicum USDA 110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bradyrhizobium sp BTAI1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brevibacillus brevis NBRC 100599</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brevundimonas subvibrioides ATCC 15264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brucella abortus A13334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brucella canis ATCC 23365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brucella melitensis bv 1 str 16M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brucella microti CCM 4915</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butyrivibrio proteoclasticus B31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor bescii DSM 6725</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor hydrothermalis 108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor kristjanssonii 177R1B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor kronotyakensis 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor lactoaceticus 6A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor obsidiensis OB47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor owensensis OL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldicellulosiruptor saccharolyticus DSM 8903</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caldivirga nitroreducens DSM 19672</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter concisus IC 167</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter fetus subsp fetus 82 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter fetus subsp fetus 82 40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter hominis ATCC BAA 381</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacter jejuni subspecies jejuni 81116</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Campylobacter lari RM2100
Candidatus Accumulibacter phosphatis clade IIA str UW 1
Candidatus Amoebophilus asiaticus 5a2
Candidatus Arthromitus sp SFB mouse Japan
Candidatus Azobacteroides pseudotrichonymphae genomovar CFP2
Candidatus Blochmannia floridanus
Candidatus Carsonella ruddii PV
Candidatus Chloracidobacterium thermophilum B
Candidatus Desulforudis audaxviator MP104C
Candidatus Hamiltonella defensa SAT Acyrthosiphon pisum
Candidatus Hodgkinia cicadica Dsem
Candidatus Korarchaeum cryptofilum OPF8
Candidatus Koribacter versatilis Ellin345
Candidatus Liberibacter asiaticus str psy62
Candidatus Methylophilus oxyfera
Candidatus Midichloria mitochondrii iricVA
Candidatus Moranella endobia PCIT
Candidatus Nitrospira defluvii
Candidatus Pelagibacter sp IMCC9063
Candidatus Phytoplasma australense
Candidatus Protochloridium amoebofila UWE25
Candidatus Puniceispirillum marinum IMCC1322
Candidatus Riesia pediculicola USDA
Candidatus Ruthia magnifica str Cm Calyptogena magnifica
Candidatus Solibacter usitatus Ellin6076
Candidatus Sulcia muelleri GWSS
Candidatus Tremblaya princeps PCIT
Candidatus Vescicomyosocius okutanii HA
Candidatus Zinderia insecticola CARI
Capnocytophaga canimorsus Cc5
Capnocytophaga ochracea DSM 7271
Carboxydothermus hydrogenoformans Z 2901
Carnobacterium sp 17 4
Catenulispora acidiphila DSM 44928
Caulobacter crescentus CB15
Caulobacter segnis ATCC 21756
Caulobacter sp K31
Cellulomonas fimii ATCC 484
Cellulomonas flavigena DSM 20109
Cellulophaga algicola DSM 14237
Cellulophaga lytica DSM 7489
Cellvibrio gilvus ATCC 13127
Cellvibrio japonicus Ueda107
Cenarchaeum symbiosum A
Chelatovorans sp BNC1
Chitinophaga pinensis DSM 2588
Chlamydia muridarum Nigg
Chlamydia trachomatis 434 Bu
Chlamydomphila abortus S26 3
Chlamydomphila caviae GPC
Chlamydomphila felis Fe C 56
Chlamydomphila pecorum E58
Chlamydomphila pneumoniae CWL029
Chlamydomphila psitacci 6BC
Chlorobaculum parvum NCIB 84928
Chlorobium chlorochromatii CaD3
Chlorobium limicola DSM 245
Chlorobium luteolum DSM 273
Chlorobium phaeobacteroides DSM 266
Chlorobium phaeovibrioides DSM 265
Chlorobium tepidum TLS
Chloroflexus aggregans DSM 9485
Chloroflexus aurantiacus J 10 fl
Chloroflexus sp Y 400 fl
Chlororherpeton thalassii ATCC 35110
Chromobacterium violaceum ATCC 12472
Chromohalobacter salexigens DSM 3043
Citrobacter koseri ATCC BAA 895
Citrobacter rodentium ICC168
Clavibacter michiganensis subsp michiganensis NCPPB 382
Clostridium acetobutylicum ATCC 824
Clostridium beijerinckii NCLB 8052
Clostridium botulinum A str ATCC 3502
Clostridium cellulolyticum H10
Clostridium cellulovorans 748B
Clostridium difficile 630
Clostridium kluveri DSM 555
Clostridium lentocellum DSM 5427
Clostridium ljungdahlii DSM 13528
Clostridium novyi NT
Clostridium perfringens ATCC 13124
Clostridium phytofermentans ISDg
Clostridium saccharolyticum WM1
Clostridium sp SY8519
Clostridium sticklandii DSM 519
Clostridium tetani E88
Clostridium thermocellum ATCC 27405
Collimonas fungivorans Ter331
Colwellia psychrerythraea 34H
SIANN: Strain Identification by Alignment to Near Neighbors

Comamonas testosteroni CNB 2
Conexibacter woesei DSM 14684
Coprothermobacter proteolyticus DSM 5265
Coralimargarita akajimensis DSM 45221
Coriobacterium glomerans PW2
Corynebacterium aurimucosum ATCC 700975
Corynebacterium diphtheriae NCTC 13129
Corynebacterium efficiens YS 314
Corynebacterium glutamicum ATCC 13032
Corynebacterium jeikeium K411
Corynebacterium kroppenstedtii DSM 44385
Corynebacterium pseudotuberculosis FRC41
Corynebacterium resistens DSM 45100
Corynebacterium ulcerans 809
Corynebacterium urealyticum DSM 7109
Corynebacterium variabile DSM 44702
Coxiella burnetii RSA 493
Croceibacter atlanticus HTCC2559
Cronobacter sakazakii ATCC BAA 894
Cronobacter turicensis 23032
Cryptobacterium curatum DSM 15641
Cupriavidus metallidurans CH34
Cupriavidus necator N 1
Cupriavidus taiwanensis LMG 19424
cyanobacterium UCYN A
Cyanobacterium sp ATCC 51142
Cyclobacterium marinus DSM 745
Cytophaga hutchinsonii ATCC 33406
Dechloromonas aromatica RCB
Deferribacter desulfuricans SSM1
Dehalococcoides ethenogenes 195
Dehalococcoides sp BAV1
Dehalogenimonas lykanthroporepellens BL DC 9
Deinococcus deserti VCD115
Deinococcus geothermals DSM 11300
Deinococcus maricopensis DSM 21211
Deinococcus proteolyticus MRP
Deinococcus radiodurans R1
Deftia acidovorans SPH 1
Deftia sp Cs1 4
Deinococcus radiodurans R1
Deftia acidovorans SPH 1
Deftia sp Cs1 4
Desulfbacca acetoxidans DSM 11109
Desulfbacca acetoxidans DSM 11109
Desulfbacterium autotrophicum HRM2
Desulfbaculabium alkenivorans AK 01
Desulfotobacterium hafniense YS1
Desulfobacca acetoxidans DSM 11109
Desulfbacterium autotrophicum HRM2
Desulfobulbus propionicus DSM 2032
Desulfococcus oleovorans Hx3
Desulfocalobacter retbaense DSM 5692
Desulfojarcibacterium baculatum DSM 4028
Desulfotalea psychrophila LSv54
Desulfotomaculum acetoxidans DSM 771
Desulfotomaculum carboxydovorans CO 1 SRB
Desulfotomaculum kuznetsovii DSM 6115
Desulfotomaculum reducens MI 1
Desulfotomaculum ruminis DSM 2154
Desulfovibrio aespoeensis Aspo 2
Desulfovibrio africanus str Walvis Bay
Desulfovibrio alaskensis G20
Desulfovibrio desulfuricans ND132
Desulfovibrio magneticus RS 1
Desulfovibrio salexigens DSM 2638
Desulfovibrio vulgaris RCH1
Desulfurispirillum indicum S5
Desulfurovibrio alkaliphilus AHT2
Desulurobacterium thermolithotrophum DSM 11699
Desulfurococcus kamchatkensis 1221n
Desulfurococcus mcosus DSM 2162
Dichelobacter nadosus VCS1703A
Dicyogolum thermophilum H 6 12
Dicyogolum turgidum DSM 6724
Dinoroseobacter shibae DSM 1247
Dyadobacter fermentans DSM 18053
E. coli E. coli ATCC 25922
E. coli E. coli ATCC 25922
Edwardsiella ictaluri 93 146
Edwardsiella tarda EIB202
Eggerthella lenta DSM 2243
Eggerthella sp YY7918
Ehrlichia canis str Jake
Ehrlichia chaffeensis str Arkansas
Ehrlichia ruminantium str Welgevonden
Eusimicrobium minutum Pei191
Enterobacter aerogenes KCTC 2190
Enterobacter asburiae LF7a
Enterobacter cloacae subsp cloacae ATCC 13047
Enterobacter sp 638
Enterococcus faecalis V583
SIANN: Strain Identification by Alignment to Near Neighbors

Erwinia amylovora ATCC 49946
Erwinia billingiae Eb661
Erwinia pyrifoliae DSM 12163
Erwinia sp Ejp617
Erwinia tasmaniensis E11 99
Eryipelothrix rhusiopathiae
Erythrobacter litoralis HTCC2594
Escherichia coli O157:H7 str Sakai
Escherichia fergusonii ATCC 35469
Ethanoligenens harbinense YUAN 3
Eubacterium eligens ATCC 27750
Eubacterium limosum KIST612
Eubacterium rectale ATCC 33656
Exiguobacterium sibiricum 255 15
Exiguobacterium sp AT1b
Ferrimonas balearica DSM 9799
Ferroglobus placidus DSM 10642
Fervidobacterium nodosum R117 B1
Fibrobacter succinogenes subsp succinogenes s85
Filifactor alocis ATCC 35896
Finegoldia magna ATCC 29328
Flavobacteriaceae bacterium 3519 10
Flavobacterium branchiophilum FL 15
Flavobacterium columnare ATCC 49512
Flavobacterium johnsoniae UW101
Flavobacterium psychrophilum JIP02 86
Flexistipes sinusarabici DSM 4947
Fluvicola taffensis DSM 16823
Francisella novicida U112
Francisella philomiragia subsp philomiragia ATCC 25017
Francisella sp TX077308
Francisella tularensis subsp holarctica LVS
Frankia alni ACN14a
Frankia sp Cci3
Frankia symbiont of Datisca glomerata
Fusobacterium nucleatum subsp nucleatum ATCC 25586
Galibacterium anatis UMN179
Gallionella capsiferriformans ES 2
Gamma proteobacterium Hdn1
Gardnerella vaginalis 409 05
Gemmatimonas aurantiaca T27
Geobacillus kaustophilus HTA426
Geobacillus sp C56 T3
Geobacillus thermodenitrificans NG80 2
Geobacillus thermoglucosidasius C56 YS93
Geobacillus thermoleovorans CCB US3 UF5
Geobacter bemidjiensis Bem
Geobacter daltonii FRC 32
Geobacter lovleyi S2
Geobacter metallireducens GS 15
Geobacter sp M18
Geobacter sulfurreducens PCA
Geobacter uraniireducens Rf4
Geodermatophilus pacificus Rf4
Geodermatophilus obscurus DSM 43160
Glaciecola nitratireducens FR1064
Glaciecola sp 4H 3 7YE 5
Gloebacter violaceus PCC 7421
Gluconacetobacter diazotrophicus PAI 5
Gluconacetobacter xylinus NBRC 3288
Gluconobacter oxydans 621H
Gordonia bronchialis DSM 43247
Gramella forsetii KT0803
Granulibacter bethesdensis CGDNIH1
Granulicella mallensis MP5ACTX8
Granulicella tundricola
Haemophilus ducreyi 35000HP
Haemophilus influenzae 10810
Haemophilus parainfluenzae T311
Haemophilus parasuis SH0165
Haemophilus somnus 129PT
Halalhalicoccus jeotgali B3
Halanaerobium hydrogeniformans
Halangium ochraceum DSM 14365
Haliscmenobacter hydrossis DSM 1100
Halaoacula hispanica ATCC 33960
Halaoacula marismortui ATCC 43049
Halobacterium sp NRC 1
Halofex volcanii DS2
Halogeometricum borinquense DSM 11551
Halomicrobium mukohataei DSM 12286
Halomonas elongata DSM 2581
halophilic archaeon DL31
Halopiger xanaduensis SH 6
Halorhodospira halophila SL1
Halorhabdus utahensis DSM 12940
Halorhabdus walsbyi C23
Halorhabdus utahensis DSM 12940
Halorhodospira halophila SL1
Halorubrum latusformis ATCC 49239
Haloterrigena turkmenica DSM 5511
Halothiobacillus neapolitanus c2
Helicobacter acinonychis str Sheeba
Helicobacter bizzozeronii CIII 1
Helicobacter felis ATCC 49179
Helicobacter hepaticus ATCC 51449
Helicobacter mustelae 12198
Helicobacter pylori 26695
SIANN: Strain Identification by Alignment to Near Neighbors

Hellobacterium modesticaldum Ice1
Herbaspirillum seropedicae SmR1
Herminimonas arsenicoxydans
Herpetosiphon aurantiacus DSM 785
Hippea maritima DSM 10411
Hirschia baltica ATCC 49814
Hydrogenobacter thermophilus TK 6
Hydrogenobaculum sp Y04AA51
Hyperthermus butylicus DSM 5456
Hyphomicrobium denitrificans ATCC 51888
Hyphomicrobium sp
Hyphomonas neptunium ATCC 15444
Idiomarina loihiensis L2TR
Ignicoccus hospitalis KIN41
Ignisphaera aggregans DSM 17230
Ilyobacter polytropus DSM 2926
Intrasporangium calvum DSM 43043
Isosphaera variabilis 225
Isosphaera pallida ATCC 43644
Jannaschia sp CCS1
Janthinobacterium sp Marseilleille
Jonesia denitrificans DSM 20603
Kangiella corensis DSM 16069
Ketogulonicigenium vulgare WSH 001
Kineococcus radiotolerans SRS30216
Kitasatospora setae KM 6054
Klebsiella oxytoca KCTC 1686
Klebsiella pneumoniae 342
Klebsiella variicola AT 22
Kocuria rhizophila DC2201
Kosmotoga olearia TBF 1951
Kribbekla flavida DSM 17836
Krokinobacter sp 4H 3 7 5
Kyrtinia tusciae DSM 2912
Kyctococcus sedentarius DSM 20547
Lacinutrix sp 5H 3 7 4
Lactobacillus acidophilus NCFM
Lactobacillus amylovorus GRL 1112
Lactobacillus brevis ATCC 367
Lactobacillus buchneri NRRL B 30929
Lactobacillus casei ATCC 334
Lactobacillus delbrueckii subsp bulgaricus ATCC 11842
Lactobacillus fermentum IFO 3956
Lactobacillus gasseri ATCC 33323
Lactobacillus helveticus DPC 4571
Lactobacillus johnsonii NCC 533
Lactobacillus kefiranofaciens ZW3
Lactobacillus plantarum subsp plantarum ST III
Lactobacillus reuteri DSM 20016
Lactobacillus rhamnosus ATCC 8530
Lactobacillus ruminis ATCC 27782
Lactobacillus sakei subsp sakei 23K
Lactobacillus salivarius UCC118
Lactobacillus sanfranciscensis TMW 11304
Lactococcus garvieae ATCC 49156
Lactococcus lactis subsp cremoris NZ9000
Laribacter hongkongensis HLHK9
Lawsonia intracellularis PHE MNI 00
Legionella pneumophila subsp pneumophila ATCC 43290
Legionella xylitolica subsp xylitolica strain CTC807
Leptospira biflexa serovar Patoc strain Patoc 1 Ames
Leptospira borgpetersenii serovar Hardjo bovis strain L550
Leptospira interrogans serovar Copenhageni strain Fiocruz L1 130
Leptothrix chalodnii SP 6
Leptotrichia buccalis C 1013 b
Leuconostoc citreum KM20
Leuconostoc gasicomitatum LMG 18811
Leuconostoc kimchii IMSNU 11154
Leuconostoc mesenteroides subsp mesenteroides ATCC 8293
Leuconostoc sp C2
Listeria innocua Clip11262
Listeria ivanovii
Listeria monocytogenes EGD e
Listeria seeligeri serovar 1 2b strain SLCC3954
Listeria welshimeri serovar 6b strain SLCC5334
Lysinibacillus sphaericus C3 41
Macrococcus caseolyticus JCSC5402
Magnetococcus marinus MC 1
Magnetobacillus magneticum AMB 1
Mahella australiensis 50 1 BON
Mannheimia succiniiproducens MBE55E
Maribacter sp HTCC2170
Maricaulis maris MCS10
Marinobacter adhaerens HP 15
Marinobacter aquaeolei VT8
Marinomonas mediterranea MMB 1
SIANN: Strain Identification by Alignment to Near Neighbors

Marinomonas posidonica IVIA Po 181
Marinomonas sp MWYL1
Marivirga tractuosa DSM 4126
Megaphera elsdenii
Melothemus ruber DSM 1279
Melothemus silvanus DSM 9446
Melissococcus plutonius ATCC 35311
Mesoplasma florum L1
Mesorhizobium ciceri biovar biserrulae WSM1271
Mesorhizobium loti MAF303099
Mesorhizobium opportunistum WSM2075
Metallosphaera cuprina Ar 4
Metallosphaera sedula DSM 5348
Methanobacterium sp AL 21
Methanobrevibacter ruminantium
M1 Methanobrevibacter smithii ATCC 35061
Methanocaldococcus fervens AG86
Methanocaldococcus infernus ME
Methanocaldococcus jannaschii DSM 2661
Methanocaldococcus sp FS406 22
Methanocaldococcus vulcanius M7
Methanocella arvoryzae MRE50
Methanocella paludicola SANAE
Methanococcoides burtonii DSM 6242
Methanococcus aeolicus Nankai 3
Methanococcus maripaludis S2
Methanococcus vannieli SB
Methanococcus voltae A3
Methanocorpusculum labreanum Z
Methanoculleus marisnigri JR1

Methanohalobium evestigatum Z 7303
Methanohalophilus mahii DSM 5219
Methanoplanus petrolearius DSM 11571
Methanopyrus kandleri AV19
Methanoregula boonei 6A8
Methanosaeta conciliae GP6
Methanosaeta thermophila PT
Methanosalum zhilinae DSM 4017
Methanosarcina acetivorans C2A
Methanosarcina Barkeri str Fusaro
Methanosarcina mazi Go1
Methanospirillum palustris E1 9c
Methanospirillum hungatei JF 1
Methanothermobacter marburgensis str Marburg
Methanothermobacter thermautotrophicus str Delta H
Methanothermococcus okinawensis IH1
Methanothermus fervidus DSM 2088
Methanotriss igneus Kol 5
Methylacidiphilum infernorum V4
Methylleibium petroleiphilum PM1
Methylobacillus flagellatus KT
Methylobacterium chloromethanicum CM4
Methylobacterium extorquens AM1
Methylobacterium nodulans ORS 2060
Methylobacterium populi BJ001
Methylobacterium radiotolerans JCM 2831
Methylobacterium sp 4 46
Methylocella silvestris BL2

Methylococcus capsulatus str Bath
Methylocystis thermautotrophicus str Delta H
Methylococcus maripaludis S2
Methylococcus voltae A3
Methanocorpusculum labreanum Z
Methanoculleus marisnigri JR1

Micrococcus luteus NCTC 2665
Microcystis aeruginosa NIES 843
Microlunatus phosphovorus NM 1
Micromonas aurantiaca ATCC 27029
Micromonas sp L5
Mobiluncus curtisi ATCC 43063
Moorella thermoacetica ATCC 39073
Moraxella catarrhalis RH4
Morulaea ruestringensis DSM 13258
Mycobacterium abscessus ATCC 19977
Mycobacterium africanum GM041182
Mycobacterium avium 104
Mycobacterium bovis AF2122
Mycobacterium canettii CIPT 140010059
Mycobacterium gilvum PYR GCK
Mycobacterium leprae TN
Mycobacterium leprae TN
Mycobacterium marinum M
Mycobacterium rhodiaceae NBB3
Mycobacterium smegmatis str MC2 155
Mycobacterium sp JDM601
Mycobacterium tuberculosis H37Rv
<table>
<thead>
<tr>
<th>Organism Name</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mycobacterium ulcerans Agy99</td>
<td></td>
</tr>
<tr>
<td>Mycobacterium vanbaalenii PYR 1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma agalactiae PG2</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma arthritidis 158L3 1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma bovis PG45</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma capricolum subsp capricolum ATCC 27343</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma conjunctivae HRC 581</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma crocodyli MP145</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma gallisepticum str Rlow</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma genitalium G37</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma haemocanis str Illinois</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma haemofelis str Langford 1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma hominis</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma hyopneumoniae 232</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma hyorhinis HUB 1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma leachii 99 014 6</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma mobile 163K</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma mycoides subsp mycoides SC str PG1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma penetrans HF 2</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma pneumoniae M129</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma pulmonis UAB CTIP</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma putrefaciens KS1</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma suis str Illinois</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma synoviae S3</td>
<td></td>
</tr>
<tr>
<td>Myxococcus fulvus HW 1</td>
<td></td>
</tr>
<tr>
<td>Myxococcus xanthus DK 1622</td>
<td></td>
</tr>
<tr>
<td>Nakamurella multipartita DSM 44233</td>
<td></td>
</tr>
<tr>
<td>Nanoarchaeum equitans Kin4 M</td>
<td></td>
</tr>
<tr>
<td>Natranaerobius thermophilus JW NM WN LF</td>
<td></td>
</tr>
<tr>
<td>Natribala magadii ATCC 43099</td>
<td></td>
</tr>
<tr>
<td>Natronomonas pharaonis DSM 2160</td>
<td></td>
</tr>
<tr>
<td>Nautilia profundicola AmH</td>
<td></td>
</tr>
<tr>
<td>Neisseria gonorrhoeae FA 1090</td>
<td></td>
</tr>
<tr>
<td>Neisseria lactamica 020 06</td>
<td></td>
</tr>
<tr>
<td>Neisseria meningitidis M01 240149</td>
<td></td>
</tr>
<tr>
<td>Neorickettsia risticii str Illinois Neorickettsia senetnsu str Miyama</td>
<td></td>
</tr>
<tr>
<td>Niastella koreensis GR20 10</td>
<td></td>
</tr>
<tr>
<td>Nitratiraptor salsuginis DSM 16511</td>
<td></td>
</tr>
<tr>
<td>Nitratiruptor sp SB155 S</td>
<td></td>
</tr>
<tr>
<td>Nitrobacter hamburgensis X14</td>
<td></td>
</tr>
<tr>
<td>Nitrobacter winogradskyi Nb 255</td>
<td></td>
</tr>
<tr>
<td>Nitrosococcus halophilus Nc4</td>
<td></td>
</tr>
<tr>
<td>Nitrosococcus oceani ATCC 19707</td>
<td></td>
</tr>
<tr>
<td>Nitrosococcus watsoni C113</td>
<td></td>
</tr>
<tr>
<td>Nitrosomonas europaea ATCC 19718</td>
<td></td>
</tr>
<tr>
<td>Nitrosomonas eutropha C91</td>
<td></td>
</tr>
<tr>
<td>Nitrosomonas sp AL212</td>
<td></td>
</tr>
<tr>
<td>Nitrospumilus maritimus SCM1</td>
<td></td>
</tr>
<tr>
<td>Nitrosospira multiformis ATCC 25196</td>
<td></td>
</tr>
<tr>
<td>Nocardia farcinica IFM 10152</td>
<td></td>
</tr>
<tr>
<td>Nocardioides sp JS614</td>
<td></td>
</tr>
<tr>
<td>Nocardiosis dasonvillei subsp dasonvillei DSM 43111</td>
<td></td>
</tr>
<tr>
<td>Nostoc azollae 0708</td>
<td></td>
</tr>
<tr>
<td>Nostoc punctiforme PCC 73102</td>
<td></td>
</tr>
<tr>
<td>Nostoc sp PCC 7120</td>
<td></td>
</tr>
<tr>
<td>Novosphingobium aromaticivorans DSM 12444</td>
<td></td>
</tr>
<tr>
<td>Novosphingobium sp PP1Y</td>
<td></td>
</tr>
<tr>
<td>Oceanithermus profundus DSM 14977</td>
<td></td>
</tr>
<tr>
<td>Oceanobacillus iheyensis HTE831</td>
<td></td>
</tr>
<tr>
<td>Ochrobacterium anthropi ATCC 49188</td>
<td></td>
</tr>
<tr>
<td>Odoribacter splanchnicus DSM 20712</td>
<td></td>
</tr>
<tr>
<td>Oenococcus oeni PSU 1</td>
<td></td>
</tr>
<tr>
<td>Oligotropha carboxidovorans OMS5</td>
<td></td>
</tr>
<tr>
<td>Olsenella uli DSM 7084</td>
<td></td>
</tr>
<tr>
<td>Onion yellows phytoplasma OY M</td>
<td></td>
</tr>
<tr>
<td>Opitutus terrae PB90 1</td>
<td></td>
</tr>
<tr>
<td>Orientia tsutsugamushi str Boryong</td>
<td></td>
</tr>
<tr>
<td>Oscillibacter valericigenes Owenweekia hongkonensis</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus mucilaginosus KNP414</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus polymyxua E681</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus sp JDR 2</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus terreae HPL 003</td>
<td></td>
</tr>
<tr>
<td>Paludibacter propionicigenes WB4</td>
<td></td>
</tr>
<tr>
<td>Pantoea ananatis LMG 20103</td>
<td></td>
</tr>
<tr>
<td>Pantoea sp At 9b</td>
<td></td>
</tr>
<tr>
<td>Pantoea vagans C9 1</td>
<td></td>
</tr>
<tr>
<td>Parabacteroides distasonis ATCC 8503</td>
<td></td>
</tr>
<tr>
<td>Parachlamydia acanthamoebae UV 7</td>
<td></td>
</tr>
<tr>
<td>Paracoccus denitrificans PD1222</td>
<td></td>
</tr>
<tr>
<td>Parvibaculum lavamentivorans DS 1</td>
<td></td>
</tr>
<tr>
<td>Parvularcula bermudensis HTCC2503</td>
<td></td>
</tr>
<tr>
<td>Pasteurella multocida subsp multocida str Pm70</td>
<td></td>
</tr>
<tr>
<td>Pectobacterium atrosepticum SCRI1043</td>
<td></td>
</tr>
<tr>
<td>Pectobacterium carotovorum subsp carotovorum PC1</td>
<td></td>
</tr>
<tr>
<td>Pectobacterium wasabiae WPP163</td>
<td></td>
</tr>
<tr>
<td>Pediococcus pentosaceus ATCC 25745</td>
<td></td>
</tr>
<tr>
<td>Pedobacter heparinus DSM 2366</td>
<td></td>
</tr>
<tr>
<td>Pedobacter saltans DSM 12145</td>
<td></td>
</tr>
<tr>
<td>Pelagibacterium halotolerans B2</td>
<td></td>
</tr>
<tr>
<td>Pelobacter carbinolicus DSM 2380</td>
<td></td>
</tr>
<tr>
<td>Pelobacter propionicus DSM 2379</td>
<td></td>
</tr>
<tr>
<td>Pelodictyon phaeoclathratiforme BU 1</td>
<td></td>
</tr>
<tr>
<td>Pelotomaculum thermopropionicum SI</td>
<td></td>
</tr>
<tr>
<td>Persephonella marina EX H1</td>
<td></td>
</tr>
<tr>
<td>Petrothoga mobilis SJ95</td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Phenylobacterium zucineum HLK1
Photobacterium profundum SS9
Photobacterium asymbiotica subsp. asymbiotica ATCC 43949
Photobacterium luminescens subsp. laumondii TTO1
Picrophilus torridus DSM 9790
Pirellula staleyi DSM 6068
Planctomyces brasiliensis DSM 5305
Planctomyces limnophilus DSM 3776
Polaromonas naphthalenivorans CJ2
Polaromonas sp JS666
Polymorphum gilvum SL003B 26A1
Polynucleobacter necessarius subsp. asymbioticus QLW P1DMWA 1
Porphyromonas asaccharolytica DSM 20707
Porphyromonas gingivalis W83
Prevotella dentiflora F0289
Prevotella melaninogenica ATCC 25845
Prevotella ruminicola 23
Prochlorococcus marinus str. MIT 9215
Propionibacterium acnes 6609
Propionibacterium freudenreichii subsp. shermanii CIRM BIA1
Prosthecochloris aestuarii DSM 271
Proteus mirabilis HI4320
Pseudoalteromonas atlantica T6c
Pseudoalteromonas haloplanktis TAC125
Pseudoalteromonas sp SM9913
Pseudogulbenkiania sp NH88B
Pseudomonas aeruginosa PAO1
Pseudomonas brassicacearum subsp. brassicacearum NFM421
Pseudomonas entomophila L48
Pseudomonas fluorescens F113
Pseudomonas fulva 12 X
Pseudomonas mendocina ymp
Pseudomonas protegens Pf 5
Pseudomonas putida F1
Pseudomonas stutzeri A1501
Pseudomonas syringae pv. phaseolicola 1448A
Pseudonocardia dioxanivorans CB1190
Pseudovibrio sp FO BEG1
Pseudoxanthomonas spadix BD a59
Pseudoxanthomonas suwonensis 11 1
Psychrobacter arcticus 273 4
Psychrobacter cryohalolentis K5
Psychrobacter sp PRw1 1
Psychromonas ingrahamii 37
Pusiillimonas sp IT7 17
Pyrococcus abyssi GE5
Pyrococcus furiosus DSM 3638
Pyrococcus horikoshii OT3
Pyrococcus sp NA2
Pyrococcus yayanosii CH1
Pyrolobus fumarii 1A
Rahnella sp Y9602
Ralstonia eutropha JMP134
Ralstonia pickettii 12D
Ralstonia solanacearum GMI1000
Ramsibacter tataoquinensis TTB310
Renibacterium salmoninarum ATCC 33209
Rhizobium etli CFN 42
Rhizobium leguminosarum bv. viciae 3841
Rhodobacter capsulatus SB 1003
Rhodobacter sphaeroides 241
Rhodococcus equi 103S
Rhodococcus erythropolis PR4
Rhodococcus jostii RHA1
Rhodococcus opacus B4
Rhodothermus marinus DSM 4252
Rickettsia africae ESF 5
Rickettsia akari str Hartford
Rickettsia belli RML369 C
Rickettsia canadensis str McKiel
Rickettsia conorii str Malish 7
Rickettsia felis URRWXCal2
Rickettsia heliobangiengensis 054
Rickettsia japonica YH
Rickettsia massiliiae MT1
Rickettsia peacockii str Rustic
Rickettsia prowazekii str Madrid E
Rickettsia rickettsii str Iowa
Rickettsia sibirica 246
Rickettsia slovaca 13 B
Rickettsia typhi str Wilmington
Rhodopirellula baltica SH 1
Rhodospirillum centenum SW
Rhodospirillum rubrum ATCC 11170
Rhodothermus marinus DSM 4252
SIANN: Strain Identification by Alignment to Near Neighbors

Roseobacter denitrificans OCh 114
Roseobacter litoralis Och 149
Rothia dentocariosa ATCC 17931
Rothia mucilaginosa DY 18
Rubrobacter xylanophilus DSM 9941
Ruegeria pomeroyi DSS 3
Ruegeria sp TM1040
Ruminococcus albus 7
Runella slithyformis DSM 19594
Saccharomonospora viridis DSM 43017
Saccharophagus degradans 2 40
Saccharopolyspora erythraea NRRL 2338
Salinibacter ruber DSM 13855
Salinispora arenicola CNS 205
Salinispora tropica CNB 440
Salmonella bongori NCTC 12419
Salmonella enterica subspp enterica serovar Enteritidis str P125109
Salmonella enterica subspp enterica serovar Enteritidis str
Salmonella boydii Sb227
Salmonella dysenteriae Sd197
Salmonella flexneri 2a str 301
Salmonella sonnei Ss046
Sideroxydans lithotrophicus ES 1
Sinorhizobium fredii NGR234
Sinorhizobium meliloti 1021
Sphingobacterium sp SYK 6
Sphingopyxis alaskensis RB2256
Sporosarcina pasteurianus ATCC BAA 2069
Streptococcus mitis B6
Streptococcus pelliculatus subsp pelliculatus
Streptococcus pneumoniae ATCC 33091
Streptococcus pyogenes ATCC 19615
Stackebrandtia nassauensis DSM 44728
Stackebrandtia nassauensis ATCC 49959
Serratia proteamaculans 568
Serratia sp AS12
Serratia symbiotica str Cinara cedri
Shewanella amazonensis SB2B
Shewanella baltica OS155
Shewanella denitrificans OS217
Shewanella frigidimarina NCIMB 400
Shewanella halifaxensis HAW EB4
Shewanella loihica PV 4
Shewanella oneidensis MR 1
Shewanella pealeana ATCC 700345
SIANN: Strain Identification by Alignment to Near Neighbors

Streptococcus pyogenes M1
Streptococcus salivarius JIM8780
Streptococcus sanguinis SK36
Streptococcus suis 052YH33
Streptococcus thermophilus CNRZ1066
Streptococcus uberis 0140J
Streptomyces avermitilis MA4680
Streptomyces bingchengensis BCW 1
Streptomyces cattleya NRRL 8057 DSM 46488
Streptomyces coelicolor A32
Streptomyces flavogriseus ATCC 33331
Streptomyces griseus subsp griseus NBRC 13350
Streptomyces scabiei 8722
Streptomyces sp SirexAA E
Streptomyces violaceusniger Tu 4113
Streptosporangium roseum DSM 43021
Sulfobacillus acidophilus DSM 10332
Sulfolobus acidocaldarius DSM 639
Sulfolobus islandicus M1425
Sulfolobus solfataricus P2
Sulfolobus tokodaii str 7
Sulfuricurvum kuijense DSM 16994
Sulfurihydrogenibium azorense Az Fu 1
Sulfurihydrogenibium sp YO3AOP1
Sulfurimonas autotrophica DSM 16294
Sulfurimonas denitrificans DSM 1251
Sulfurospirillum deleyanum DSM 6946
Sulfovum sp NBC37 1
Syntrophobacter thermophilum IAM 14863
Synechocystis sp PCC 6803
Syntrophobacter fumaroxidans MPOB
Syntrophobatus glycolicus DSM 8271
Syntrophomonas wolfei subsp wolfei str Goettingen
Syntrophothermus lipocalidus DSM 12680
Syntrophus aciditrophicus SB
Tannerella forsythia ATCC 43037
Taylorella equigenitalis MCE9
Tepidanaerobacter acetatoxydans Re1
Teredinibacter tumerae T7901
Terriglobus saanensis SP1PR4
Tetragenococcus halophilus
Thauera sp MZ1T
Thermabacterium maritum DSM 12885
Thermanaerobacter acidominovorans DSM 6589
Thermococcus potens JR
Thermoanaerobacter brockii subsp finnii Ako 1
Thermoanaerobacter italicus Ab9
Thermoanaerobacterium thermosaccharolyticum DSM 571
Thermoanaerobacterium xylanolyticum LX 11
Thermoanaerobacter mathranii subsp mathranii str A3
Thermoanaerobacter pseudelathanolicus ATCC 33223
Thermoanaerobacter sp X513
Thermoanaerobacter tengcongensis MB4
Thermoanaerobacter wiegelli R1881
Thermobacterium terrenum ATCC BAA 798
Thermobifida fusca YX
Thermobispora bispora DSM 43833
Thermococcus barophilus MP
Thermococcus gammatolerans EJ3
Thermococcus kodakakensis KOD1
Thermococcus onnurineus NA1
Thermococcus sibiricus MM739
Thermococcus sp 4557
Thermococcus albus DSM 14484
Thermodesulfatator indicus DSM 15286
Thermodesulfobacterium sp OPB45
Thermodesulfobium narugense DSM 14796
Thermodesulfovibrio yellowstonii DSM 11347
Thermophilum pendens Hrk 5
Thermomicrobium roseum DSM 5159
Thermomonospora curvata DSM 43183
Thermoplasma acidophilum DSM 1728
Thermoplasma volcanium GSS1
Thermoproteus uzoniensis 76820
Thermosediminibacter oceanicum DSM 16646
Thermosipho africanus TCF52B
Thermosipho melanesiensis BI429
Thermosphaera aggregans DSM 11486
Thermosynechococcus elongatus BP 1
Thermotoga lettingae TMO
Thermotoga maritima MSB8
Thermotoga naphthophila RKU 10
Thermotoga neapolitana DSM 4359
Thermotoga petrophila RKU 1
Thermotoga sp RQ2
Thermotoga thermarum DSM 5069
Thermovibrio ammonificans HB 1
Thermovirga lienii DSM 17291
SIANN: Strain Identification by Alignment to Near Neighbors

Thermus scotoductus SA 01
Thermus sp CCB US3 UF1
Thermus thermophilus HB27
Thioalkalimicrobium cyclicum ALM1
Thioalkalivibrio sp K90mix
Thioalkalivibrio sulfidophilus HL EbGr7
Thiobacillus denitrificans ATCC 25259
Thiomicrospira crunogena XCL 2
Thiomonas intermedia K12
Tolumonas auensis DSM 9187
Treponema azotonutricium ZAS 9
Treponema brennaborens DSM 12168
Treponema denticola ATCC 35405
Treponema pallidum subsp pallidum str Nichols
Treponema paralucisuniculi Cuniculi A
Treponema primitia ZAS 2
Treponema succinifaciens DSM 2489
Trichodesmium erythraeum IMS101
Tropheryma whipplei str Twist
Truepera radiovictrix DSM 17093
Tsukamurella paurometabola DSM 20162
Ureaplasma parvum serovar 3 str ATCC 27815
Ureaplasma urealyticum serovar 10 str ATCC 33699
Variovorax paradoxus S110
Veillonella parvula DSM 2008
Verminephrobacter eiseniae EF01 2
Verrucosispora maris AB 18 032
Vibrio anguillarum 775
Vibrio cholerae O1 biovar El Tor str N16961
Vibrio fischeri ES114
Vibrio furnissii NCTC 11218
Vibrio harveyi ATCC BAA 1116
Vibrio parahaemolyticus RIMD
2210633
Vibrio sp EJY3
Vibrio splendidus LGP32
Vibrio vulnificus CMCP6
Vulcanisaeta distriubuta DSM 14429
Vulcanisaeta moukovskia 768 28
Waddlia chondrophila WSU 86 1044
Weeksellia virosa DSM 16922
Weissella koreensis KACC 15510
Wigglesworthia glossinidia endosymbiont of Glossina brevipalpis
Wolbachia endosymbiont of Culex quinquefasciatus Pel
Wolbachia sp wRi
Wolinella succinogenes DSM 1740
Xanthobacter autotrophicus Py2
Xanthomonas albineans GPE PC73
Xanthomonas axonopodis pv citri str 306
Xanthomonas campestris pv campestris str 8004
Xanthomonas oryzae pv oryzae KACC 10331
Xenorhabdus bovienii SS 2004
Xenorhabdus nematophila ATCC 19061
Xylanimonas cellulosilytica DSM 15894
Xylella fastidiosa 9a5c
Yersinia enterocolitica subsp enterocolitica 8081
Yersinia pestis CO92
Yersinia pseudotuberculosis IP 32953
Zobella galactanivorans
Zunongwangia profunda SM A87
Zymomonas mobilis subsp mobilis ATCC 10988