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.

![Figure 1](image)

**Figure 1.** A) For a group of strains belonging to two different species, some regions may be unique to each species (region 1), while other regions may be unique to strains within each species (regions 2 and 3). B) A set of reads are aligned to these genomes, and the ones that align in a species- or strain-specific manner are identified by the combination of genomes to which they align. In this example, Strain B of Species I is the organism identified.

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
SIANN: Strain Identification by Alignment to Near Neighbors

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>Bacillus anthracis</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>Bacillus cereus</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>Hanta virus</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></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clostridium botulinum</td>
<td>19%</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>
</tr>
<tr>
<td>Clostridium difficile</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>Listeria fleischmannii</td>
<td></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>
</tr>
<tr>
<td>Listeria monocytogenes</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>1.2%</td>
<td>4.7%</td>
<td>19%</td>
<td>76%</td>
<td>0.074%</td>
<td>0.30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaccinia virus</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>Yersinia enterocolitica</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>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yersinia pestis</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>19%</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.
SIANN: Strain Identification by Alignment to Near Neighbors

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 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.

Appendices

Appendix 1: Target Pathogen Database

**Arenaviridae**
- Arenavirus
  - Golden-Gate-virus
  - Arenavirus Lujo-virus
  - Flexal-virus segment-L
  - New-world-arenaviruses Allpahuayo-virus
  - New-world-arenaviruses Chapare-virus
  - New-world-arenaviruses Guanarito-virus
  - New-world-arenaviruses Junin-virus
  - New-world-arenaviruses Machupo-virus
  - New-world-arenaviruses Sabia-virus
  - New-world-arenaviruses Tacaribe-virus
  - New-world-arenaviruses Whitewater-Arroyo-virus
  - Old-world-arenaviruses Ippy-virus
  - Old-world-arenaviruses Lassa-virus
  - Old-world-arenaviruses Mopeia-virus-AN20410

**Asfarviridae**
- African-swine-fever-virus Benin-971-pathogenic-isolate
- African-Swine-Fever-Virus

**Bacillaceae**
- Anthracis A2012 Bant 02 1
- Anthracis Ames Ancestor
- Anthracis Sterne
- cereus 03BB102
- cereus AH187
- cereus AH820
- cereus ATCC 14579
- cereus B4264
- cereus F65185
- cytotoxicus NVH 391-98
- mycoides DSM 2048
- mycoides Rock1-4
- thuringiensis BMB171
- thuringiensis B1407
- thuringiensis HD-771
- thuringiensis serovar chinensis CT-43
- thuringiensis serovar konukian 97-27
- abortus A13334
- cettia B1 94
- cetti M13 05 1 supercont1 22
- melitensis ATCC 23457
- melitensis bv1 str16M
- ovis ATCC 25840
- suis 1330
- suis ATCC 23445
- Akabane-virus segment-M
- Hantavirus
- Andes-virus
- Hantavirus
- Dobrava-Belgrade-virus-strain-DOBV-Ano-Poroia-Alfl9-1999
- Hantavirus
- Puumala-virus
- Hantavirus
- Seoul-virus-strain-Seoul-80-39-clone-1
- Hantavirus
- Sin Nombre-virus
- Hantavirus
- Crimean-Congo-hemorrhagic-fever-virus
- Hantavirus
- Nairobi-virus
- Dugbe-virus
- Phlebovirus Rift-Valley-fever-virus
- cenocepacia HI2424
- cenocepacia J2315
- cenocepacia MC0-3
- cepacia GG4
- gladioli BSR3
- glumae BGR1
- mallei
- mallei ATCC 10399
- mallei NCTC 10247
- mallei SAVP1
- multivorans ATCC 17616
- oklahomensis C6786
- oklahomensis EO147
- pseudomallei 1026b
- pseudomallei 1106a
- pseudomallei 1710b
- pseudomallei 668
- pseudomallei BPC006
- pseudomallei K96243
- pyrocinia CH-67
- thailandensis ATCC 700388
- thailandensis E264
- thailandensis MSMB121
- coli JV20
- fetus subsp fetus 82-40
- jejuni RM1221
SIANN: Strain Identification by Alignment to Near Neighbors

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 symbosum WAL-14163
Clostridium symbosum 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
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 SJ108 111
Coxiella burnetii RSA 493
Diplorickettsia 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-dIvoire-ebolavirus
Filoviridae
Ebolavirus
Ebola-virus-Mayinga-Zaire
Filoviridae
Ebolavirus
Reston-ebolavirus
Filoviridae
Ebolavirus
Sudan-ebolavirus
Filoviridae
Marburgvirus
Lake-Victoria-marburgvirus-Musoke
Filoviridae
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
SIANN: Strain Identification by Alignment to Near Neighbors

Flaviviridae Langat-virus
Flaviviridae Louping-ill-virus
Flaviviridae Murray-Valley-encephalitis-virus
Flaviviridae Omsk-hemorrhagic-fever-virus
Flaviviridae Powassan-virus
Flaviviridae St-Louis-encephalitis-virus
Flaviviridae Tick-borne-encephalitis-virus
Flaviviridae Usutu-virus
Flaviviridae West-Nile-virus
Flaviviridae Yellow-fever-virus
Francisella cf novicida Fx1
Francisella noatunensis subsp orientalis str Toba 04
Francisella novicida U112
Francisella philomiragia subsp philomiragia ATCC 25015
Francisella philomiragia subsp philomiragia ATCC 25017
Francisella tularensis subsp holarctica F92
Francisella tularensis subsp mediasiatica FSC147
Herpesviridae Alcelaphine-herpesvirus 1
Herpesviridae Acelaphine-herpesvirus 1
Listeria monocytogenes FSL R2-561
Listeria monocytogenes FSL N1-067
Listeria seeligeri FSL 12b str SLCC3954
Listeria welschimeri serovar 6b str SLCC5334
Paramyxoviridae Avulavirus Newcastle-disease-virus-B1
Paramyxoviridae Henipavirus Hendra-virus
Paramyxoviridae Henipavirus Nipah-virus
Paramyxoviridae Menangle-virus
Paramyxoviridae Morbillivirus Measles-virus
Paramyxoviridae Peste-des-petits-ruminants-virus
Paramyxoviridae Respirovirus Human-parainfluenza-virus-1
Paramyxoviridae Respirovirus Human-parainfluenza-virus-3
Paramyxoviridae Respirovirus Sendai-virus
Paramyxoviridae Rinderpest-virus strain-Kabete-O
Paramyxoviridae Rubulavirus Human-parainfluenza-virus-2
Paramyxoviridae Rubulavirus Human-parainfluenza-virus-4a
Paramyxoviridae Rubulavirus Mumps-virus
Picornaviridae Foot-and-mouth-disease-virus -type-O
Picornaviridae Swine-vesicular-disease-virus strain-HK70
Picornaviridae Swine-vesicular-disease-virus strain-NET192
Poxviridae Avipoxvirus Fowlpox-virus
Poxviridae Crocodylipsexvirus Nile-crocodylipsexvirus
Poxviridae Goatpox-virus Pellar
Poxviridae Leporipoxvirus Myxoma-virus
Poxviridae Lumpy-skin-disease-virus NI-2490
Poxviridae Molluscipoxvirus Molluscum-contagiosum-virus-subtype-1
Poxviridae Orthopoxvirus Camelpox-virus
Poxviridae Orthopoxvirus Cowpox-virus
Poxviridae OrthopoxvirusECTROMELIA- virus
Poxviridae Orthopoxvirus Monkeypox-virus Zaire-96-16
Poxviridae Orthopoxvirus Taterapox-virus
Poxviridae Orthopoxvirus Vaccinia-virus
Poxviridae Orthopoxvirus Voriola-virus
Poxviridae Sheeppox-virus 17077-99
Poxviridae Suipoxvirus
Poxviridae Swinepox-virus
Puccinia graminis f sp tritici
CRL 75-36-700-3
Ralstonia pickettii 12D
Ralstonia pickettii 12J
Ralstonia solanacearum CFBP2957
Ralstonia solanacearum CMR15
Ralstonia solanacearum GMI1000
Rathayibacter toxicus DSM 7488
Reoviridae African-horsesickness-virus segment-10
Rhabdoviridae Vesicular-stomatitis-Indiana-virus
Rhabdoviridae Vesicular-stomatitis-virus strain-NJ2075212NM
Rickettsia bellii OSU 85-389
Rickettsia conorii str Malish 7
Rickettsia prowazekii str Breinl
Rickettsia prowazekii str RpGvF24
Rickettsia rickettsii str Arizona
Rickettsia typhi str B9991CWPP
Rickettsiella grylli gcontig 634
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 Baranah-Forest-virus
Togaviridae Chikungunya-virus
Togaviridae EEEV-complex Eastern-equine-encephalitis-virus
Togaviridae Rubivirus Rubella-virus
Togaviridae SFV-complex O-nyong-nyong-virus
Togaviridae Venezuelean-equine-encephalitis-virus
Togaviridae WEEV-complex Sindbis-virus
Togaviridae Western-equine-encephalomyelitis-virus
Xanthomonas albilineans GPE PC73
Xanthomonas axonopodis Xac29-1
Xanthomonas oryzae pv oryzae PXO99A
Xanthomonas oryzae pv oryzae MAFF 311018
Xanthomonas oryzae pv oryzae PC73
Xanthomonas vasicola pv vasculorum 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 Pestoides F
Yersinia pseudotuberculosis IP 31758
Yersinia pseudotuberculosis IP 32953
Yersinia pseudotuberculosis PB1
Yersinia pseudotuberculosis YP1
Yersinia ruckeri ATCC 29473
Xanthomonas oryzae pv oryzae KACC 10331
Xanthomonas oryzae pv oryzae MAFF 311018
Xanthomonas oryzae pv oryzae PXO99A
Xanthomonas vasicola pv vasculorum 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 Pestoides F
Yersinia pseudotuberculosis IP 31758
Yersinia pseudotuberculosis IP 32953
Yersinia pseudotuberculosis PB1
Yersinia pseudotuberculosis YP1
Yersinia ruckeri ATCC 29473
Xanthomonas oryzae pv oryzae KACC 10331
## Appendix 2: Viral Database

<table>
<thead>
<tr>
<th>Virus/Family</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abaca bunchy top virus DNA-C</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abaca bunchy top virus DNA-M</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abaca bunchy top virus DNA-N</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abaca bunchy top virus DNA-R</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abaca bunchy top virus DNA-S</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abaca bunchy top virus segment 2</td>
<td>Abaca bunchy top virus</td>
</tr>
<tr>
<td>Abalone shriveling syndrome-</td>
<td>Abalone shriveling syndrome-related virus</td>
</tr>
<tr>
<td>Abutilon Brazil virus DNA A</td>
<td>Abutilon Brazil virus</td>
</tr>
<tr>
<td>Abutilon Brazil virus DNA B</td>
<td>Abutilon Brazil virus</td>
</tr>
<tr>
<td>Abutilon mosaic virus DNA A</td>
<td>Abutilon mosaic virus</td>
</tr>
<tr>
<td>Abutilon mosaic virus DNA B</td>
<td>Abutilon mosaic virus</td>
</tr>
<tr>
<td>Acanthocystis turfacea</td>
<td>Acanthocystis turfacea</td>
</tr>
<tr>
<td>Chlorella virus 1</td>
<td>Chlorella virus</td>
</tr>
<tr>
<td>Acheta domesticus</td>
<td>Acheta domesticus</td>
</tr>
<tr>
<td>densovirus</td>
<td>Densovirus</td>
</tr>
<tr>
<td>Acholeplasma phage L2</td>
<td>Acholeplasma phage</td>
</tr>
<tr>
<td>Acholeplasma phage MV-L1</td>
<td>Acholeplasma phage</td>
</tr>
<tr>
<td>Acidianus bottle-shaped virus</td>
<td>Acidianus bottle-shaped virus</td>
</tr>
<tr>
<td>Acidianus filamentous virus 1</td>
<td>Acidianus filamentous virus 1</td>
</tr>
<tr>
<td>Acidianus filamentous virus 2</td>
<td>Acidianus filamentous virus 2</td>
</tr>
<tr>
<td>Acidianus filamentous virus 3</td>
<td>Acidianus filamentous virus 3</td>
</tr>
<tr>
<td>Acidianus filamentous virus 6</td>
<td>Acidianus filamentous virus 6</td>
</tr>
<tr>
<td>Acidianus filamentous virus 7</td>
<td>Acidianus filamentous virus 7</td>
</tr>
<tr>
<td>Acidianus filamentous virus 8</td>
<td>Acidianus filamentous virus 8</td>
</tr>
<tr>
<td>Acidianus filamentous virus 9</td>
<td>Acidianus filamentous virus 9</td>
</tr>
<tr>
<td>Acidianus rod-shaped virus 1</td>
<td>Acidianus rod-shaped virus 1</td>
</tr>
<tr>
<td>Acidianus spindle-shaped virus 1</td>
<td>Acidianus spindle-shaped virus 1</td>
</tr>
<tr>
<td>Acidianus two-tailed virus</td>
<td>Acidianus two-tailed virus</td>
</tr>
<tr>
<td>Actinomycetes phage Av-1</td>
<td>Actinomycetes phage</td>
</tr>
<tr>
<td>Actinoplanes phage phiAsp2</td>
<td>Actinoplanes phage</td>
</tr>
<tr>
<td>Acyrhospilon pismum bacteria phage</td>
<td>Acyrhospilon pismum mafia phage</td>
</tr>
<tr>
<td>APSE-1</td>
<td>APSE-1</td>
</tr>
<tr>
<td>Adeno-associated virus-1</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-2</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-3</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-4</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-5</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-7</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adeno-associated virus-8</td>
<td>Adeno-associated virus</td>
</tr>
<tr>
<td>Adoxophytes orana</td>
<td>Adoxophytes orana</td>
</tr>
<tr>
<td>granulovirus</td>
<td>Granulovirus</td>
</tr>
<tr>
<td>Adoxophytes orana</td>
<td>Adoxophytes orana</td>
</tr>
<tr>
<td>nucleopolyhedrovirus</td>
<td>Nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Aedes aegypti densovirus</td>
<td>Aedes aegypti densovirus</td>
</tr>
<tr>
<td>Aedes albopictus densovirus</td>
<td>Aedes albopictus densovirus</td>
</tr>
<tr>
<td>Aedes taeniorhynchus iridescent</td>
<td>Aedes taeniorhynchus iridescent</td>
</tr>
<tr>
<td>virus</td>
<td>Virus</td>
</tr>
<tr>
<td>Aeromonas phage 25</td>
<td>Aeromonas phage</td>
</tr>
<tr>
<td>Aeromonas phage 44RR2.8t</td>
<td>Aeromonas phage</td>
</tr>
<tr>
<td>Aeromonas phage Aeh1</td>
<td>Aeromonas phage</td>
</tr>
<tr>
<td>Aeromonas phage phiO18P</td>
<td>Aeromonas phage</td>
</tr>
<tr>
<td>african cassava mosaic virus DNA A</td>
<td>African cassava mosaic virus DNA A</td>
</tr>
<tr>
<td>african cassava mosaic virus DNA B</td>
<td>African cassava mosaic virus DNA B</td>
</tr>
<tr>
<td>African green monkey polyomavirus</td>
<td>African green monkey polyomavirus</td>
</tr>
<tr>
<td>African swine fever virus</td>
<td>African swine fever virus</td>
</tr>
<tr>
<td>Ageratum enation virus</td>
<td>Ageratum enation virus</td>
</tr>
<tr>
<td>Ageratum leaf Cameroon betasatellite</td>
<td>Ageratum leaf Cameroon betasatellite</td>
</tr>
<tr>
<td>Ageratum leaf curl virus-G52</td>
<td>Ageratum leaf curl virus-G52</td>
</tr>
<tr>
<td>Ageratum yellow vein China virus-associated DNA beta</td>
<td>Ageratum yellow vein China virus-associated DNA beta</td>
</tr>
<tr>
<td>Ageratum yellow vein Chinavirus</td>
<td>Ageratum yellow vein Chinavirus</td>
</tr>
<tr>
<td>Ageratum yellow vein Hualian virus-TaiwanHsinchutom2003 DNA A</td>
<td>Ageratum yellow vein Hualian virus-TaiwanHsinchutom2003 DNA A</td>
</tr>
<tr>
<td>Ageratum yellow vein Sri Lanka virus segment A</td>
<td>Ageratum yellow vein Sri Lanka virus segment A</td>
</tr>
<tr>
<td>Ageratum yellow vein Taiwan virus</td>
<td>Ageratum yellow vein Taiwan virus</td>
</tr>
<tr>
<td>Ageratum yellow vein virus-associated DNA beta</td>
<td>Ageratum yellow vein virus-associated DNA beta</td>
</tr>
<tr>
<td>Agrotis epsilon multiple nucleopolyhedrovirus</td>
<td>Agrotis epsilon multiple nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Agrotis segetum granulovirus</td>
<td>Agrotis segetum granulovirus</td>
</tr>
<tr>
<td>Agrotis segetum</td>
<td>Agrotis segetum</td>
</tr>
<tr>
<td>nucleopolyhedrovirus</td>
<td>Nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Alcelaphine herpesvirus 1</td>
<td>Alcelaphine herpesvirus</td>
</tr>
<tr>
<td>Aleutian mink disease virus</td>
<td>Aleutian mink disease virus</td>
</tr>
<tr>
<td>Allamanda leaf curl virus DNA-A</td>
<td>Allamanda leaf curl virus DNA-A</td>
</tr>
<tr>
<td>Alternanthera yellow vein virus DNA-A</td>
<td>Alternanthera yellow vein virus DNA-A</td>
</tr>
<tr>
<td>Alternanthera yellow vein virus satellite DNA beta</td>
<td>Alternanthera yellow vein virus satellite DNA beta</td>
</tr>
<tr>
<td>Ambystoma tigrinum virus</td>
<td>Ambystoma tigrinum virus</td>
</tr>
<tr>
<td>Anmacta moorei</td>
<td>Anmacta moorei</td>
</tr>
<tr>
<td>entomopoxovirus L</td>
<td>Entomopoxovirus</td>
</tr>
<tr>
<td>Anguillid herpesvirus 1</td>
<td>Anguillid herpesvirus</td>
</tr>
<tr>
<td>Anopheles gambiae</td>
<td>Anopheles gambiae</td>
</tr>
<tr>
<td>densovirus</td>
<td>Densovirus</td>
</tr>
<tr>
<td>Antheraea permyn</td>
<td>Antheraea permyn</td>
</tr>
<tr>
<td>nucleopolyhedrovirus</td>
<td>Nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Anticarsia gemmatalis</td>
<td>Anticarsia gemmatalis</td>
</tr>
<tr>
<td>nucleopolyhedrovirus</td>
<td>Nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Archeaeal BJ1 virus</td>
<td>Archeaeal BJ1 virus</td>
</tr>
<tr>
<td>Ateline herpesvirus 3</td>
<td>Ateline herpesvirus</td>
</tr>
<tr>
<td>Autographa californica</td>
<td>Autographa californica</td>
</tr>
<tr>
<td>nucleopolyhedrovirus</td>
<td>Nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Avian adeno-associated virus</td>
<td>Avian adeno-associated virus</td>
</tr>
<tr>
<td>ATCC VR-865</td>
<td>ATCC VR-865</td>
</tr>
<tr>
<td>Avian adeno-associated virus strain DA-1</td>
<td>Avian adeno-associated virus strain DA-1</td>
</tr>
<tr>
<td>Avian endogenous retrovirus</td>
<td>Avian endogenous retrovirus</td>
</tr>
<tr>
<td>EAV-HP</td>
<td>EAV-HP</td>
</tr>
<tr>
<td>Azospirillum phage Cd</td>
<td>Azospirillum phage</td>
</tr>
<tr>
<td>Bacillus phage 0305phi8-36</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage AP50</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage B103</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage Bam35c</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage BCJA1c</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage Cherry</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage Fah</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage GA-1</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage Gamma</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage GIL16c</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage IEBH</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage phi105</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage phi29</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage SP8c2</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage SPO1</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage SPP1</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage TP21-L</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus phage WBeta</td>
<td>Bacillus phage</td>
</tr>
<tr>
<td>Bacillus prophage phBC6A51</td>
<td>Bacillus prophage</td>
</tr>
<tr>
<td>Bacillus prophage phBC6A52</td>
<td>Bacillus prophage</td>
</tr>
<tr>
<td>Bacillus virus 1</td>
<td>Bacillus virus</td>
</tr>
<tr>
<td>Bacteriophage Aaphi23</td>
<td>Bacteriophage Aaphi23</td>
</tr>
<tr>
<td>Bacteriophage APSE-2</td>
<td>Bacteriophage APSE-2</td>
</tr>
<tr>
<td>Bacteriophage PSA</td>
<td>Bacteriophage PSA</td>
</tr>
<tr>
<td>Bacteriophage R832</td>
<td>Bacteriophage R832</td>
</tr>
<tr>
<td>Bacteroides phage B40-8</td>
<td>Bacteroides phage B40-8</td>
</tr>
</tbody>
</table>
### SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Virus/Phage Name</th>
<th>Associated DNA/Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banana bunchy top virus DNA C</td>
<td>Begomovirus-associated DNA II</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA M</td>
<td>Begomovirus-associated DNA-III</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA N</td>
<td>Bettongia penicillata papilomavirus 1</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA R</td>
<td>Bhendi yellow vein Bhubhaneswar virus DNA-A</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA S</td>
<td>Bhendi yellow vein Delhi virus 2004New Delhi DNA-A</td>
</tr>
<tr>
<td>Banana bunchy top virus DNA U3</td>
<td>Bhendi yellow vein mosaic virus-associated DNA beta</td>
</tr>
<tr>
<td>Banana streak GF virus</td>
<td>Bhendi yellow vein mosaic virus</td>
</tr>
<tr>
<td>Banana streak Mysore virus</td>
<td>Bitter gourd leaf curl disease-associated DNA beta</td>
</tr>
<tr>
<td>Banana streak OL virus</td>
<td>BK polyomavirus</td>
</tr>
<tr>
<td>Banana streak virus genome</td>
<td>Blainvillea yellow spot virus DNA-A</td>
</tr>
<tr>
<td>Banana streak virus strain Acuminata Vietnam</td>
<td>Blainvillea yellow spot virus DNA-B</td>
</tr>
<tr>
<td>Bandicoot papillomatosis carcinoma virus type 1</td>
<td>Blattella germanica densovirus</td>
</tr>
<tr>
<td>Bandicoot papillomatosis carcinoma virus type 2</td>
<td>Blueberry red ringspot virus</td>
</tr>
<tr>
<td>Bat adeno-associated virus YNM</td>
<td>Bocavirus gorillaGBov12009</td>
</tr>
<tr>
<td>Bdellovibrio phage phiMH2K</td>
<td>Bombyx mandarina nucleopolyhedrovirus</td>
</tr>
<tr>
<td>Beak and feather disease virus</td>
<td>Bombyx mori densovirus 5</td>
</tr>
<tr>
<td>Bean calico mosaic virus DNA A</td>
<td>Bombyx mori NPV</td>
</tr>
<tr>
<td>Bean calico mosaic virus DNA B</td>
<td>Bordetella phage BIP-1</td>
</tr>
<tr>
<td>Bean dwarf mosaic virus DNA A</td>
<td>Bordetella phage BMP-1</td>
</tr>
<tr>
<td>Bean dwarf mosaic virus DNA B</td>
<td>Bordetella phage BPP-1</td>
</tr>
<tr>
<td>Bean golden mosaic virus DNA A</td>
<td>Bougainvillea spectabilis chlorotic vein-banding virus</td>
</tr>
<tr>
<td>Bean golden mosaic virus DNA B</td>
<td>Bovine aden-associated virus</td>
</tr>
<tr>
<td>Bean golden yellow mosaic virus DNA A</td>
<td>Bovine adenovirus A</td>
</tr>
<tr>
<td>Bean golden yellow mosaic virus DNA B</td>
<td>Bovine adenovirus B</td>
</tr>
<tr>
<td>Bean yellow dwarf virus putative genes V1</td>
<td>Bovine adenovirus D</td>
</tr>
<tr>
<td>Beet curly top Iran virus-K</td>
<td>Bovine ephemeral fever virus</td>
</tr>
<tr>
<td>Beet curly top virus-California Logan</td>
<td>Bovine foamy virus</td>
</tr>
<tr>
<td>Beet mild curly top virus-Worland4</td>
<td>Bovine herpesvirus 1</td>
</tr>
<tr>
<td>Beet severe curly top virus-Cfh</td>
<td>Bovine herpesvirus 4 long unique region</td>
</tr>
<tr>
<td></td>
<td>Bovine herpesvirus 5</td>
</tr>
<tr>
<td></td>
<td>Bovine papillomavirus-1</td>
</tr>
<tr>
<td></td>
<td>Bovine papillomavirus 3</td>
</tr>
<tr>
<td></td>
<td>Bovine papular stomatitis virus</td>
</tr>
<tr>
<td></td>
<td>Bovine parvovirus 2</td>
</tr>
<tr>
<td></td>
<td>Bovine Parvovirus</td>
</tr>
<tr>
<td></td>
<td>Bovine polyomavirus</td>
</tr>
<tr>
<td></td>
<td>Burkholderia ambifaria phage BcepF1</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage Bcep 176</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage Bcep1</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage Bcep22</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage Bcep43</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage Bcep781</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage BcepB1A</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage BcepC6B</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage BcepGommr</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage BcepIL02</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage BcepMu</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiE25</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiE12</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phil026b</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiE12-2</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiE202</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiE255</td>
</tr>
<tr>
<td></td>
<td>Burkholderia phage phiS237</td>
</tr>
<tr>
<td></td>
<td>Cabbage leaf curl virus DNA A</td>
</tr>
<tr>
<td></td>
<td>Cabbage leaf curl virus DNA B</td>
</tr>
<tr>
<td></td>
<td>Cacao swollen shoot virus</td>
</tr>
<tr>
<td></td>
<td>California sea lion anellovirus</td>
</tr>
<tr>
<td></td>
<td>California sea lion polyomavirus 1</td>
</tr>
<tr>
<td></td>
<td>Callitrichine herpesvirus 3 strain CJ0149</td>
</tr>
<tr>
<td></td>
<td>Camelpox virus</td>
</tr>
<tr>
<td></td>
<td>Campoletis sonorenisis ichnovirus chromosome segment W</td>
</tr>
<tr>
<td></td>
<td>Campoletis sonorenisis ichnovirus segment B</td>
</tr>
<tr>
<td></td>
<td>Campoletis sonorenisis ichnovirus segment C</td>
</tr>
<tr>
<td></td>
<td>Campoletis sonorenisis ichnovirus segment D</td>
</tr>
<tr>
<td></td>
<td>Campoletis sonorenisis ichnovirus segment E</td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Campoletis sonorensis ichnovirus segment F
Campoletis sonorensis ichnovirus segment G2
Campoletis sonorensis ichnovirus segment G
Campoletis sonorensis ichnovirus segment H
Campoletis sonorensis ichnovirus segment I2
Campoletis sonorensis ichnovirus segment I
Campoletis sonorensis ichnovirus segment J
Campoletis sonorensis ichnovirus segment L
Campoletis sonorensis ichnovirus segment M
Campoletis sonorensis ichnovirus segment N
Campoletis sonorensis ichnovirus segment O1
Campoletis sonorensis ichnovirus segment P
Campoletis sonorensis ichnovirus segment Q
Campoletis sonorensis ichnovirus segment T
Campoletis sonorensis ichnovirus segment U
Campoletis sonorensis ichnovirus segment V
Campoletis sonorensis ichnovirus segment Z
Campoletis sonorensis ichnovirus superhelical segment A
Campoletis sonorensis ichnovirus superhelical segment aprime
Capra hircus papillomavirus type 1
Cardiospernum yellow leaf curl virus satellite DNA beta
Caretta caretta papillomavirus 1
Carnation etched ring virus
Casphalia extranea densovirus
Cassava vein mosaic virus
Cauliflower mosaic virus
Cercopithecine herpesvirus 2
Cercopithecine herpesvirus 5 strain 2715
Cercopithecine herpesvirus 9
Cestrum yellow leaf curling virus
Chaetoceros salsugineum DNA virus
Chayote yellow mosaic virus
Chicken anemia virus
Chickpea chlorotic dwarf virus
Chili leaf curl disease associated sequence virion
Chili leaf curl Multan alphastellite
Chili leaf curl virus
Chino del tomate virus DNA A
Chino del tomate virus DNA B
Chlamydia phage 3
Chlamydia phage 4
Chlamydia phage Chp1
Chlamydia phage Chp2
Chlamydia phage PhiCPG1
Chloris striate mosaic virus
Choristoneura fumiferana DEF MNPV
Choristoneura fumiferana MNPV
Choristoneura occidentalis granulovirus
Chrysodeixis chalcites nucleopolyhedrovirus
Circovirus-like genome BBC-A
Circovirus-like genome CB-A
Circovirus-like genome CB-B
Circovirus-like genome RW-A
Circovirus-like genome RW-B
Circovirus-like genome RW-C
Circovirus-like genome RW-D
Circovirus-like genome RW-E
Circovirus-like genome SAR-A
Circovirus-like genome SAR-B
Citrus psorosis virus RNA1
Citrus psorosis virus RNA2
Citrus psorosis virus RNA3
Citrus yellow mosaic virus
Clanis bilineata nucleopolyhedrosis virus
Clavibacter phage CMP1
Clerodendron yellow mosaic virus
Clerodendrum golden mosaic virus
China virus DNA A
Clerodendrum golden mosaic virus DNA A
Clerodendrum golden mosaic virus DNA-B
Clostridium phage 39-O
Clostridium phage c-st
Clostridium phage phi3626
Clostridium phage phiC2
Clostridium phage phi CD119
Clostridium phage phiCD27
Clostridium phage phiCTP1
Coconut foliar decay virus
Columbid circovirus
Commelina yellow mottle virus
Corchorus golden mosaic virus DNA-A
Corchorus golden mosaic virus DNA-B
Corchorus yellow spot virus DNA A
Corchorus yellow spot virus DNA B
Corchorus yellow vein virus-Hoa Binh DNA A
Corchorus yellow vein virus-Hoa Binh DNA B
Corynebacterium phage BFK20
Corynebacterium phage P1201
Cotesia congregata bracovirus segment Circle10
Cotesia congregata bracovirus segment Circle11
Cotesia congregata
bracovirus segment Circle12
Cotesia congregata
bracovirus segment Circle13
Cotesia congregata
bracovirus segment Circle14
Cotesia congregata
bracovirus segment Circle15
Cotesia congregata
bracovirus segment Circle17
Cotesia congregata
bracovirus segment Circle19
Cotesia congregata
bracovirus segment circle1
Cotesia congregata
bracovirus segment Circle20
Cotesia congregata
bracovirus segment Circle21
Cotesia congregata
bracovirus segment Circle22
Cotesia congregata
bracovirus segment Circle23
Cotesia congregata
bracovirus segment Circle25
Cotesia congregata
bracovirus segment Circle26
Cotesia congregata
bracovirus segment circle2
Cotesia congregata
bracovirus segment Circle30
Cotesia congregata
bracovirus segment Circle31
Cotesia congregata
bracovirus segment Circle32
Cotesia congregata
bracovirus segment Circle33
Cotesia congregata
bracovirus segment Circle35
Cotesia congregata
bracovirus segment Circle36
Cotesia congregata
bracovirus segment circle3
Cotesia congregata
bracovirus segment Circle4
Cotesia congregata
bracovirus segment Circle5
Cotesia congregata
bracovirus segment Circle9
Cotesia congregata virus segment Circle18
Cotesia congregata virus segment Circle6
Cotesia congregata virus segment Circle7
Cotesia congregata virus segment Circle8
Cotton leaf crumple geminivirus DNA B
Cotton leaf crumple virus DNA A
Cotton leaf curl Alabad virus
Cotton leaf curl Bangalore virus-associated DNA beta
Cotton leaf curl Bangalore virus segment A
Cotton leaf curl Burewala alphasatellite
Cotton leaf curl Burewala betasatellite
Cotton leaf curl Burewala virus-IndiaVehari2004
Cotton leaf curl Gezira alphasatellite
Cotton leaf curl Gezira betasatellite
Cotton leaf curl Gezira extrachromosomal
Cotton leaf curl Gezira virus
Cotton leaf curl Kokhran virus
Cotton leaf curl Multan Virus
Cotton leaf curl Multan virus satellite DNA beta
Cotton leaf curl Multan virus satellite U36-1
Cotton leaf curl Rajasthan virus segment A
Cotton leaf curl virus-associated DNA beta
Cottontail rabbit papillomavirus
Cowpea severe leaf curl-associated DNA beta
Cowpox virus
Craseocephalum yellow vein virus-Jinghong
Croton yellow vein mosaic alphassatellite
Croton yellow vein mosaic virus
Croton yellow vein mosaic virus satellite DNA beta
Croton yellow vein virus
Crow polyomavirus
Cryptoplebia leucotreta granulovirus
Cucurbita yellow vein virus-associated DNA beta
Cucurbit leaf crumple virus DNA A
Cucurbit leaf crumple virus DNA B
Culex nigripalpus NPV
Culex pipiens densovirus
Cyanophage PSS2
Cyanophage Syn5
Cycad leaf necrosis virus
Cydia pomonella granulovirus
Cyprinid herpesvirus 3
Deer papillomavirus
Deerpox virus W-1170-84
Deerpox virus W-848-83
Deftia phage phiW-14
Dendrolimus punctatus densovirus
Desmodium leaf distortion virus DNA A
Desmodium leaf distortion virus DNA B
Diadromus pulchellus
ascovirus 4a
Diatraea saccharalis densovirus
Diciiptera yellow mottle virus DNA A
Diciiptera yellow mottle virus DNA B
Digitaria streak virus
Dioscorea bacilliform virus
Dolichos yellow mosaic virus
Dracaea mottle virus
Duck adenovirus A
Duck circovirus
Duck hepatitis B virus
East African cassava mosaic virus
Cameroon virus DNA A
East African cassava mosaic virus
Cameroon virus DNA B
East African cassava mosaic virus
Kenya virus DNA A
East African cassava mosaic virus
Kenya virus DNA B
SIANN: Strain Identification by Alignment to Near Neighbors

East African cassava mosaic virus DNA A
East African cassava mosaic virus DNA B
East African cassava mosaic Zanzibar virus DNA-A
East African cassava mosaic Zanzibar virus DNA B
Ectocarpus siliculosus virus 1
Ecotropis obliqua NPV
Emiliania huxleyi virus 86
Emilia yellow vein virus-associated DNA beta
Emilia yellow vein virus-Fz1
Enterobacteria phage 13a
Enterobacteria phage 933W
Enterobacteria phage alpha3
Enterobacteria phage BA14
Enterobacteria phage BP4795
Enterobacteria phage EcoDS1
Enterobacteria phage EPS7
Enterobacteria phage epsilon15
Enterobacteria phage ES18
Enterobacteria phage Felix 01
Enterobacteria phage Fels-2
Enterobacteria phage G4
Enterobacteria phage HK022
Enterobacteria phage HK620
Enterobacteria phage HK97
Enterobacteria phage I2-2
Enterobacteria phage ID18
Enterobacteria phage ID2
MoscowID2001
Enterobacteria phage If1
Enterobacteria phage Ike
Enterobacteria phage IM08
Enterobacteria phage JK06
Enterobacteria phage JS10
Enterobacteria phage JS98
Enterobacteria phage JSE
Enterobacteria phage K1-5
Enterobacteria phage K1E
Enterobacteria phage K1F
Enterobacteria phage lambda
Enterobacteria phage LKA1
Enterobacteria phage M13
Enterobacteria phage Min27
Enterobacteria phage Mu
Enterobacteria phage N15
Enterobacteria phage N4
Enterobacteria phage P1
Enterobacteria phage P22
Enterobacteria phage P2 virus
Enterobacteria phage P4
Enterobacteria phage Phi1
Enterobacteria phage phiEco32
Enterobacteria phage phiEcoM-GJ1
Enterobacteria phage phiP27
Enterobacteria phage phiX174 sensu lato
Enterobacteria phage PRD1
Enterobacteria phage PsP3
Enterobacteria phage RB14
Enterobacteria phage RB16
Enterobacteria phage RB43
Enterobacteria phage RB49
Enterobacteria phage RB51
Enterobacteria phage RB69
Enterobacteria phage RTP
Enterobacteria phage St6
Enterobacteria phage SV
Enterobacteria phage SP6
Enterobacteria phage SSL2009a
Enterobacteria phage ST104
Enterobacteria phage SI-1
Enterobacteria phage ST64T
Enterobacteria phage T1
Enterobacteria phage T3
Enterobacteria phage T4
Enterobacteria phage T5
Enterobacteria phage T7
Enterobacteria phage TLS
Enterobacteria phage VT2-Sakai
Enterobacteria phage WA13
Enterobacteria phage WV8
Enterobacteria phage YYZ2008
Enterococcus faecalis 62 chromosome
Enterococcus faecalis 62 plasmid EF62pA
Enterococcus faecalis 62 plasmid EF62pB
Enterococcus faecalis 62 plasmid EF62pC
Enterococcus phage EF62phi
Enterococcus phage EFAP1
Enterococcus phage phiEF11
Enterococcus phage phiEF24C
Enterococcus phage phiFL1A
Enterococcus phage phiFL2A
Enterococcus phage phiFL3A
Enterococcus phage phiFL4A
Enzootic nasal tumour virus of goats
Epiphyas postvittana NPV
Equine papillomavirus 2
Eragrostis curvula streak virus
Eragrostis streak virus
Erectites yellow mosaic virus DNA-A
Erectites yellow mosaic virus satellite DNA beta
Erethizon dorsatum papillomavirus type 1
Erinaceus europaeus papillomavirus
Erwinia amylovora phage Era103
Erwinia phage phiEa21-4
Escherichia coli bacteriophage rv5
Escherichia phage D108
Escherichia phage phiV10
Eupatorium vein clearing virus
Eupatorium yellow vein associated DNA beta
Eupatorium yellow vein virus
Euphorbia leaf curl virus DNA A
Euphorbia yellow mosaic virus DNA A
Euphorbia yellow mosaic virus DNA B
Euproctis pseudoconspersa nucleopolyhedrovirus
European elk papillomavirus
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 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
Finch circovirus
Finch polyomavirus
Flavobacterium phage 11b
Fowl adenovirus A
Fowl adenovirus D
Fowlpox virus
Francolinus leucoscepus papillomavirus
Fringilla coelebs papillomavirus
Frog adenovirus 1
Frog virus 3
Gallidia meliella 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-B1
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
SIANN: Strain Identification by Alignment to Near Neighbors

Glypta fumiferanae ichnovirus segment B41
Glypta fumiferanae ichnovirus segment B42
Glypta fumiferanae ichnovirus segment B43
Glypta fumiferanae ichnovirus segment B44
Glypta fumiferanae ichnovirus segment B45
Glypta fumiferanae ichnovirus segment B46
Glypta fumiferanae ichnovirus segment B47
Glypta fumiferanae ichnovirus segment B48
Glypta fumiferanae ichnovirus segment B49
Glypta fumiferanae ichnovirus segment B50
Glypta fumiferanae ichnovirus segment B51
Glypta fumiferanae ichnovirus segment B52
Glypta fumiferanae ichnovirus segment B53
Glypta fumiferanae ichnovirus segment B54
Glypta fumiferanae ichnovirus segment B55
Glypta fumiferanae ichnovirus segment B56
Glypta fumiferanae ichnovirus segment B57
Glypta fumiferanae ichnovirus segment B58
Glypta fumiferanae ichnovirus segment B59
Glypta fumiferanae ichnovirus segment B60
Glypta fumiferanae ichnovirus segment B61
Glypta fumiferanae ichnovirus segment B62
Glypta fumiferanae ichnovirus segment B63
Glypta fumiferanae ichnovirus segment B64
Glypta fumiferanae ichnovirus segment B65
Glypta fumiferanae ichnovirus segment B66
Glypta fumiferanae ichnovirus segment B67
Glypta fumiferanae ichnovirus segment B68
Glypta fumiferanae ichnovirus segment B69
Glypta fumiferanae ichnovirus segment B70
Glypta fumiferanae ichnovirus segment B71
Glypta fumiferanae ichnovirus segment B72
Glypta fumiferanae ichnovirus segment B73
Glypta fumiferanae ichnovirus segment B74
Glypta fumiferanae ichnovirus segment B75
Glypta fumiferanae ichnovirus segment B76
Glypta fumiferanae ichnovirus segment B77
Glypta fumiferanae ichnovirus segment B78
Glypta fumiferanae ichnovirus segment B79
Glypta fumiferanae ichnovirus segment B80
Glypta fumiferanae ichnovirus segment B81
Glypta fumiferanae ichnovirus segment B82
Glypta fumiferanae ichnovirus segment B83
Glypta fumiferanae ichnovirus segment B84
Glypta fumiferanae ichnovirus segment B85
Glypta fumiferanae ichnovirus segment B86
Glypta fumiferanae ichnovirus segment B87
Glypta fumiferanae ichnovirus segment B88
Glypta fumiferanae ichnovirus segment B89
Glypta fumiferanae ichnovirus segment B90
Glypta fumiferanae ichnovirus segment B91
Glypta fumiferanae ichnovirus segment B92
Glypta fumiferanae ichnovirus segment B93
Glypta fumiferanae ichnovirus segment B94
Glypta fumiferanae ichnovirus segment B95
Glypta fumiferanae ichnovirus segment B96
Glypta fumiferanae ichnovirus segment B97
Glypta fumiferanae ichnovirus segment B98
Glypta fumiferanae ichnovirus segment B99
Glypta fumiferanae ichnovirus segment C1
Glypta fumiferanae ichnovirus segment C2
Glypta fumiferanae ichnovirus segment C3
Glypta fumiferanae ichnovirus segment C4
Glypta fumiferanae ichnovirus segment C5
Glypta fumiferanae ichnovirus segment C6
Glypta fumiferanae ichnovirus segment C7
Glypta fumiferanae ichnovirus segment C8
Glypta fumiferanae ichnovirus segment C9
Glypta fumiferanae ichnovirus segment D1
Glypta fumiferanae ichnovirus segment D2
Glypta fumiferanae ichnovirus segment D3
Glypta fumiferanae ichnovirus segment D4
Glypta fumiferanae ichnovirus segment D5
Glypta fumiferanae ichnovirus segment D6
Glypta fumiferanae ichnovirus segment D7
Glypta fumiferanae ichnovirus segment E1
Goatpox virus Pellor
Goose circovirus
Goose hemorrhagic polyomavirus
Goose parvovirus
Gossypium darwinii symptomless alphasatellite DNA-alpha
Gossypium darwinii symptomless virus DNA-A
Gossypium davidsonii symptomless alphasatellite DNA-alpha-B
Gossypium mustilinum symptomless alphasatellite DNA-alpha-B
Gossypium punctatum mild leaf curl virus DNA A
Gossypium punctatum mild leaf curl virus DNA B
Ground squirrel hepatitis virus
Gryllus bimaculatus nudivirus
Gull circovirus
Haemophilus phage HP1
Haemophilus phage HP2
Haloarcula hispanica pleomorphic virus 1
Haloarcula phage SH1
Halomonas phage phiHAP-1
SIANN: Strain Identification by Alignment to Near Neighbors

- Halorubrum phage HF2
- Halorubrum pleomorphic virus 1
- Halovirus HF1
- Helicoverpa armigera granulovirus
- Helicoverpa armigera multiple nucleopolyhedrovirus
- Helicoverpa armigera-NPV
- Helicoverpa armigera NPV NNg1
- Helicoverpa armigera nucleopolyhedrovirus G4
- Helicoverpa zea SNPV
- Heliolthi virescens ascovirus 3e
- Hepatitis B virus
- Heron hepatitis B virus
- His1 virus
- His2 virus
- Hollyhock leaf crumple virus
- Honeysuckle yellow vein beta-JapanFukui2001
- Honeysuckle yellow vein mosaic 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 S4
- Human adenovirus AHuman 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 5 strain Merlin
- Human bocavirus 6
- Human bocavirus 7
- Human bocavirus 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 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 papillomavirus type 19
- 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-B1
- Hyposoter fugitivus ichnovirus segment B2
- 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

Lactobacillus phage Lc-Nu
Lactobacillus phage LL-H
Lactobacillus phage LP65
Lactobacillus phage Lrm1
Lactobacillus phage Lv-1
Lactobacillus phage phiAT3
Lactobacillus phage phiG1e
Lactobacillus phage phiJL-1
Lactobacillus prophage Lj928
Lactobacillus prophage Lj965
Lactobacillus prophage phiad
Lactococcus phage 1706
Lactococcus phage 4268
Lactococcus phage 712
Lactococcus prophage asccphi28
Lactococcus phage bILB29
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 QS4
Lactococcus phage r11
Lactococcus phage sk1
Lactococcus phage TP901-1
Lactococcus phage Tuc2009
Lactococcus phage ul36
Lactococcus prophage bl285
Lactococcus prophage bl286
Lactococcus prophage bl309
Lactococcus prophage bl310
Lactococcus prophage bl311
Lactococcus prophage bl312
Lamium leaf distortion associated virus
Leucania separata nuclear polyhedrosis virus
Leucas zeylanica yellow vein virus satellite DNA beta
SIANN: Strain Identification by Alignment to Near Neighbors

Lindernia anagallis yellow vein virus DNA-A
Lindernia anagallis yellow vein virus satellite DNA beta
Listeria phage A006
Listeria phage A118
Listeria phage A500
Listeria phage A511
Listeria phage B025
Listeria phage B054
Listeria phage P35
Listeria phage P40
Listonella phage phiHSIC
Loofa yellow mosaic virus DNA A
Loofa yellow mosaic virus DNA B
Lucky bamboo bacilliform virus
Ludwigia leaf distortion betasatellite
IndiaAmadalavalasaHibiscus2007
Ludwigia yellow vein virus-associated DNA beta
Ludwigia yellow vein virus DNA-A
Luffa begomovirus associated DNA beta
Luffa puckering and leaf distortion-associated DNA beta
Lulli virus
Lumpy skin disease virus NI-2490
Lymnaeia dispar MNPV
Lymnaeia xylina MNPV
Lymphocystis disease virus 1
Lymphocystis disease virus-isolate China
Maccacine herpesvirus 1
Maccacine herpesvirus 3
Maccacine herpesvirus 4
Maccacine herpesvirus 5
Macaque simian foamy virus
Macroptilium golden mosaic virus-
JamaicaWissadulaAugust
Town DNA B
Macroptilium mosaic Puerto Rico virus DNA A
Macroptilium mosaic Puerto Rico virus DNA B
Macroptilium yellow mosaic Florida virus DNA A
Macroptilium yellow mosaic Florida virus DNA B
Macroptilium yellow mosaic virus DNA A
Macroptilium yellow mosaic virus DNA B
Maize streak virus-ASouth Africa
Malacira yellow vein mosaic virus-associated satellite DNA beta
Mal de Rio Cuarto virus segment 9
Malvastrum leaf curl Guangdong virus
Malvastrum leaf curl virus-associated defective DNA beta
Malvastrum leaf curl virus-G87
Malvastrum yellow mosaic virus-associated DNA 1
Malvastrum yellow mosaic virus DNA-A
Malvastrum yellow mosaic virus satellite DNA beta
Malvastrum yellow vein Baoshan virus DNA-A
Malvastrum yellow vein virus satellite DNA beta
Malvastrum yellow vein Yunnan-virus
Malvastrum yellow vein Yunnan virus satellite DNA beta
Mamestra configurata NPV-A
Mamestra configurata NPV-B
Mannheimiaphage phiMHaA1
Maruca vitrata MNPV
Mastomys coucha papillomavirus 2
Mastomys natalensis papillomavirus
Melanoplus sanguinipes entomopoxvirus
Meleagrid herpesvirus 1
Melon chlorotic leaf curl virus DNA A
Melon chlorotic mosaic virus-associated alphasatellite
Melon chlorotic mosaic virus DNA-A
Melon chlorotic mosaic virus DNA-B
Merkel cell polyomavirus
Merremia mosaic virus DNA A
Merremia mosaic virus DNA B
Mesta yellow vein mosaic Bahiraich virus-IndiaBahiraich2007 DNA A
Mesta yellow vein mosaic virus-associated DNA beta
Mesta yellow vein mosaic virus DNA A
Methanobacterium phage psiM2
Methanothermabacter prophage psiM100
Microbacterium phage Min1
Microcystis phage Ma-LMM01
Microplitis demolitor bracovirus segment A
Microplitis demolitor bracovirus segment B
Microplitis demolitor bracovirus segment C
Microplitis demolitor bracovirus segment D
Microplitis demolitor bracovirus segment E
Microplitis demolitor bracovirus segment F
Microplitis demolitor bracovirus segment G
Microplitis demolitor bracovirus segment H
Microplitis demolitor bracovirus segment I
Microplitis demolitor bracovirus segment J
Microplitis demolitor bracovirus segment K
<table>
<thead>
<tr>
<th>Virus/Phage Name</th>
<th>Virus/Phage Name</th>
<th>Virus/Phage Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microplitis demolitor bracovirus segment L</td>
<td>Microplitis demolitor bracovirus segment M</td>
<td>Microplitis demolitor bracovirus segment N</td>
</tr>
<tr>
<td>Milk vetch dwarf virus segment 10</td>
<td>Milk vetch dwarf virus segment 11</td>
<td>Milk vetch dwarf virus segment 1</td>
</tr>
<tr>
<td>Milk vetch dwarf virus segment 2</td>
<td>Milk vetch dwarf virus segment 3</td>
<td>Milk vetch dwarf virus segment 4</td>
</tr>
<tr>
<td>Milk vetch dwarf virus segment 5</td>
<td>Milk vetch dwarf virus segment 6</td>
<td>Milk vetch dwarf virus segment 7</td>
</tr>
<tr>
<td>Milk vetch dwarf virus segment 8</td>
<td>Milk vetch dwarf virus segment 9</td>
<td>Mimosa yellow leaf curl virus-associated DNA 1</td>
</tr>
<tr>
<td>Mimosa yellow leaf curl virus DNA A</td>
<td>Mimosa yellow leaf curl virus DNA-B</td>
<td>Miscanthus streak virus-91</td>
</tr>
<tr>
<td>Molluscum contagiosum virus subtype 1</td>
<td>Minute virus of mice</td>
<td>Mirabilis mosaic virus</td>
</tr>
<tr>
<td>Monkeypox virus Zaire-96-I-16</td>
<td>Morganella phage MmP1</td>
<td>Morskankask virus-91</td>
</tr>
<tr>
<td>Mouse mammary tumor virus</td>
<td>Mouse parovirus 1</td>
<td>Mouse parovirus 2</td>
</tr>
<tr>
<td>Mouse parovirus 2</td>
<td>Mouse parovirus 3</td>
<td>Mouse parovirus 4</td>
</tr>
<tr>
<td>Mouse parovirus 5</td>
<td>Mulard duck circovirus</td>
<td>Mungbean yellow mosaic virus DNA-A</td>
</tr>
<tr>
<td>Mycobacterium phage Che9c</td>
<td>Mycobacterium phage Che9d</td>
<td>Mycobacterium phage Cjw1</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
<tr>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
<td>Mycobacterium phage</td>
</tr>
</tbody>
</table>
## SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Virus/Microorganism</th>
<th>Associated with:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mycoplasma phage</strong></td>
<td>MAV1, P1, phiMFV1, Mycoplasma phage VM-2008, Myxococcus phage Mx8, Myxoma virus</td>
</tr>
<tr>
<td><strong>Mycoplasma phage</strong></td>
<td>MAV1, P1, phiMFV1, Mycoplasma phage VM-2008, Myxococcus phage Mx8, Myxoma virus</td>
</tr>
<tr>
<td><strong>Peptide</strong></td>
<td>Solon, Spud, TM4, Troll4, Tweety, U2, Wildcat</td>
</tr>
<tr>
<td><strong>Pelargonium vein banding</strong></td>
<td>virus DNA B</td>
</tr>
<tr>
<td><strong>Okra yellow mosaic</strong></td>
<td>Mexico virus DNA B</td>
</tr>
<tr>
<td><strong>Okra yellow vein</strong></td>
<td>disease associated sequence virion</td>
</tr>
<tr>
<td><strong>Okra yellow mosaic</strong></td>
<td>virus Old World harvest mouse papillomavirus</td>
</tr>
<tr>
<td><strong>Orchid</strong></td>
<td>leucostigma NPV</td>
</tr>
<tr>
<td><strong>Orchid</strong></td>
<td>pseudosugata MNPV</td>
</tr>
<tr>
<td><strong>Oryctes rhinoceros</strong></td>
<td>virus</td>
</tr>
<tr>
<td><strong>Ostreococcus</strong></td>
<td>tauri virus 1</td>
</tr>
<tr>
<td><strong>Ostreococcus</strong></td>
<td>virus OsV5</td>
</tr>
<tr>
<td><strong>Oxine adenovirus</strong></td>
<td>A</td>
</tr>
<tr>
<td><strong>Oxine adenovirus</strong></td>
<td>D</td>
</tr>
<tr>
<td><strong>Oxine herpesvirus</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Oxine papillomavirus</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Panama streak virus</strong></td>
<td>Karino Heberling</td>
</tr>
<tr>
<td><strong>Panay leaf curl China</strong></td>
<td>virus G8</td>
</tr>
<tr>
<td><strong>Panay leaf curl China</strong></td>
<td>virus satellite DNA beta</td>
</tr>
<tr>
<td><strong>Panay leaf curl Guandong</strong></td>
<td>virus-GD2 DNA A</td>
</tr>
<tr>
<td><strong>Panay leaf curl virus</strong></td>
<td>-associated DNA beta</td>
</tr>
<tr>
<td><strong>Panay leaf curl-virus</strong></td>
<td>Papiine herpesvirus 2</td>
</tr>
<tr>
<td><strong>Paramaecium</strong></td>
<td>bursaria</td>
</tr>
<tr>
<td><strong>Chlorella</strong></td>
<td>virus 1</td>
</tr>
<tr>
<td><strong>Chlorella</strong></td>
<td>virus AR158</td>
</tr>
<tr>
<td><strong>Chlorella</strong></td>
<td>virus FR483</td>
</tr>
<tr>
<td><strong>Chlorella</strong></td>
<td>virus NY2A</td>
</tr>
<tr>
<td><strong>Parvovirus</strong></td>
<td>H1</td>
</tr>
<tr>
<td><strong>Passionfruit</strong></td>
<td>severe leaf distortion virus DNA-A</td>
</tr>
<tr>
<td><strong>Passionfruit</strong></td>
<td>severe leaf distortion virus DNA-B</td>
</tr>
<tr>
<td><strong>Pentatoma</strong></td>
<td>leaf curl virus</td>
</tr>
<tr>
<td><strong>Pedilanthus</strong></td>
<td>leaf curl virus</td>
</tr>
<tr>
<td><strong>Pedilanthus</strong></td>
<td>leaf curl virus</td>
</tr>
<tr>
<td><strong>PakistanMultan2004</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Pelargonium</strong></td>
<td>vein banding virus</td>
</tr>
<tr>
<td><strong>Penaeus</strong></td>
<td>merguiensis densovirus</td>
</tr>
<tr>
<td><strong>Penaeus</strong></td>
<td>monodon hepatopancreatic parvovirus</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>curly top virus</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>golden mosaic virus DNA A</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>golden mosaic virus DNA B</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>huasteco yellow vein virus DNA A</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>huasteco yellow vein virus DNA B</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>leaf curl Bangladesh virus segment A component</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>leaf curl virus DNA-A</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>leaf curl virus satellite DNA beta</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>leaf curl Yunnan virus satellite DNA beta</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>leaf curl Yunnan virus-YN323</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>yellow dwarf virus-New Mexico</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>yellow leaf curl Indonesia virus DNA-A</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>yellow leaf curl Indonesia virus DNA-B</td>
</tr>
<tr>
<td><strong>Pepper</strong></td>
<td>yellow vein Mali virus</td>
</tr>
<tr>
<td><strong>Periplaneta</strong></td>
<td>fuliginosa densovirus</td>
</tr>
<tr>
<td><strong>Petunia</strong></td>
<td>vein clearing virus</td>
</tr>
<tr>
<td><strong>Phage</strong></td>
<td>cdtl</td>
</tr>
<tr>
<td><strong>Phage</strong></td>
<td>Gifsy-1</td>
</tr>
<tr>
<td><strong>Phage</strong></td>
<td>Gifsy-2</td>
</tr>
<tr>
<td><strong>Phage</strong></td>
<td>phiJL001</td>
</tr>
<tr>
<td><strong>Phocoena</strong></td>
<td>spinipinnis papillomavirus</td>
</tr>
<tr>
<td><strong>Phoridium</strong></td>
<td>phage Pf-WMP3</td>
</tr>
<tr>
<td><strong>Phthorimaea</strong></td>
<td>operculella granulovirus</td>
</tr>
<tr>
<td><strong>Pieris</strong></td>
<td>rapae granulovirus</td>
</tr>
<tr>
<td><strong>Planaria</strong></td>
<td>asexual strain-specific virus-like element type 1 large DNA segment</td>
</tr>
<tr>
<td><strong>Planaria</strong></td>
<td>asexual strain-specific virus-like element type 1 small DNA segment</td>
</tr>
</tbody>
</table>
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
Pseudaelia 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 PaP3
Pseudomonas phage PB1
Pseudomonas phage Pf1
Pseudomonas phage Pi3
Pseudomonas phage phi-2
Pseudomonas phage phiCTX
Pseudomonas phage phiK77
Pseudomonas phage phiKMV
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 RS1B
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 phage 16-3
Rhodothrumus phage 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 phage SiO1
Roseophage DSS3P2
Roseophage EE36P1
Ross goose hepatitis B virus
Roussetus 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
<table>
<thead>
<tr>
<th>Virus/Phage Name</th>
<th>DNA/RNA Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Sida</em> golden mosaic Costa Rica virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic Costa Rica virus</td>
<td>DNA B</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic Florida virus-Malvastrum</td>
<td>DNA-A</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic Florida virus-Malvastrum</td>
<td>DNA-B</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic Honduras virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic Honduras virus</td>
<td>DNA B</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> golden mosaic virus</td>
<td>DNA-B</td>
</tr>
<tr>
<td><em>Sida</em> golden mottle virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> golden mottle virus</td>
<td>DNA-B</td>
</tr>
<tr>
<td><em>Sida</em> leaf curl virus-associated DNA 1</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> leaf curl virus-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> leaf curl virus</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> leaf curl virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> mircrantha mosaic virus</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> mircrantha mosaic virus</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> mircrantha mosaic virus segment A</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> mircrantha mosaic virus segment B</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> mosaic Sinaloa virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> mosaic Sinaloa virus</td>
<td>DNA B</td>
</tr>
<tr>
<td><em>Sida</em> mottle virus</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow mosaic virus-China-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow mosaic virus</td>
<td>DNA beta</td>
</tr>
<tr>
<td><em>Sida</em> yellow mosaic virus Yucatan virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> yellow mosaic Yucatan virus</td>
<td>DNA B</td>
</tr>
<tr>
<td><em>Sida</em> yellow vein disease associated DNA 1</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow vein Madurai virus</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow vein Vietnam virus-associated DNA 1</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow vein Vietnam virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> yellow vein Vietnam virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Sida</em> yellow vein virus</td>
<td>DNA A</td>
</tr>
<tr>
<td><em>Sida</em> yellow vein virus</td>
<td>DNA B</td>
</tr>
<tr>
<td><em>Sida</em> yellow vein virus satellite DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Siegesbeckia</em> yellow vein Guangxi virus</td>
<td></td>
</tr>
<tr>
<td><em>Siegesbeckia</em> yellow vein virus-GD13-associated DNA beta</td>
<td></td>
</tr>
<tr>
<td><em>Siegesbeckia</em> yellow vein virus GD13</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> adenovirus 1</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> adenovirus 3</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> foamy virus 3</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> immunodeficiency virus SIV-mnd 2</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> retrovirus 4</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> T-cell lymphotropic virus 6</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> T-lymphotropic virus 1</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> T-lymphotropic virus 3</td>
<td></td>
</tr>
<tr>
<td><em>Simian</em> virus 40</td>
<td></td>
</tr>
<tr>
<td>Singapore grouper iridovirus</td>
<td></td>
</tr>
<tr>
<td>Sinorhizobium phage PBC5</td>
<td></td>
</tr>
<tr>
<td>Small anellovirus 1</td>
<td></td>
</tr>
<tr>
<td>Small anellovirus 2</td>
<td></td>
</tr>
<tr>
<td>Snake parvovirus 1</td>
<td></td>
</tr>
<tr>
<td>Snow goose hepatitis B virus</td>
<td></td>
</tr>
<tr>
<td>Sodalis phage phiSG1</td>
<td></td>
</tr>
<tr>
<td>Sodalis phage SO-1</td>
<td></td>
</tr>
<tr>
<td>South African cassava mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>South African cassava mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Soybean chlorotic blotch virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Soybean chlorotic blotch virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Soybean chlorotic mottle virus</td>
<td></td>
</tr>
<tr>
<td>Soybean crinkle leaf virus</td>
<td></td>
</tr>
<tr>
<td>Soybean mild mottle virus</td>
<td></td>
</tr>
<tr>
<td>Spilanthes yellow vein virus DNA-A</td>
<td></td>
</tr>
<tr>
<td>Spinach curly top virus</td>
<td></td>
</tr>
<tr>
<td>Spiroplasma kunkelii virus SkV1 CR2-3x</td>
<td></td>
</tr>
<tr>
<td>Spiroplasma phage 1-C74</td>
<td></td>
</tr>
<tr>
<td>Spiroplasma phage 1-R8A2B</td>
<td></td>
</tr>
<tr>
<td>Spiroplasma phage 4</td>
<td></td>
</tr>
<tr>
<td>Spiroplasma phage SVTS2</td>
<td></td>
</tr>
<tr>
<td>Spodoptera exigua MNPV</td>
<td></td>
</tr>
<tr>
<td>Spodoptera frugiperda ascovirus 1a</td>
<td></td>
</tr>
<tr>
<td>Spodoptera frugiperda MNPV virus</td>
<td></td>
</tr>
<tr>
<td>Spodoptera litura granulovirus</td>
<td></td>
</tr>
<tr>
<td>Spodoptera litura NPV</td>
<td></td>
</tr>
<tr>
<td>Spodoptera litura nucleopolyhedrovirus II</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl China virus-B DNA A</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl China virus-B DNA B</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl Philippines virus segment A</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl Philippines virus segment B</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl virus A component DNA</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl virus B component DNA</td>
<td></td>
</tr>
<tr>
<td>Squash leaf curl Yunnan virus</td>
<td></td>
</tr>
<tr>
<td>Squash mild leaf curl virus-Imperial Valley DNA A</td>
<td></td>
</tr>
<tr>
<td>Squash mild leaf curl virus-Imperial Valley DNA B</td>
<td></td>
</tr>
<tr>
<td>Squash yellow mild mottle virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Squirrel monkey polyomavirus</td>
<td></td>
</tr>
<tr>
<td>Sri Lankan cassava mosaic virus DNA A</td>
<td></td>
</tr>
<tr>
<td>Sri Lankan cassava mosaic virus DNA B</td>
<td></td>
</tr>
<tr>
<td>Stachytarpheta leaf curl virus</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 11</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 187</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 2638A</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 29</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 37</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 3A</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 42e</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 44AHJD</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 47</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 52A</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 53</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 55</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 66</td>
<td></td>
</tr>
<tr>
<td>Staphylococcus phage 69</td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

<table>
<thead>
<tr>
<th>Staphylococcus phage 71</th>
<th>Staphylococcus phage phi 12</th>
<th>Staphylococcus prophage MM1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staphylococcus phage 77</td>
<td>Staphylococcus phage phi 13</td>
<td>Streptococcus pyogenes phage 315.3</td>
</tr>
<tr>
<td>Staphylococcus phage 80alpha</td>
<td>Staphylococcus phage phiN315</td>
<td>Streptomyces phage mu16</td>
</tr>
<tr>
<td>Staphylococcus phage 85</td>
<td>Staphylococcus phage phiPV83</td>
<td>Streptomyces phage phiB1T1</td>
</tr>
<tr>
<td>Staphylococcus phage 88</td>
<td>Staphylococcus phage PVL</td>
<td>Streptomyces phage phiC31</td>
</tr>
<tr>
<td>Staphylococcus phage 92</td>
<td>Staphylococcus phage phiSASD1</td>
<td>Streptomyces phage phiVWB</td>
</tr>
<tr>
<td>Staphylococcus phage 96</td>
<td>Staphylococcus phage phiLA2</td>
<td>Stx1-converting phage</td>
</tr>
<tr>
<td>Staphylococcus phage CNPH82</td>
<td>Staphylococcus phage phiLA2</td>
<td>Stx2-converting phage 1717</td>
</tr>
<tr>
<td>Staphylococcus phage EW</td>
<td>Staphylococcus phage phiLA2</td>
<td>Stx2-converting phage 86</td>
</tr>
<tr>
<td>Staphylococcus phage G1</td>
<td>Staphylococcus phage phiLA2</td>
<td>Stx2-converting phage II</td>
</tr>
<tr>
<td>Staphylococcus phage K</td>
<td>Streptococcus prophage phi3396</td>
<td>Stx2-converting phage</td>
</tr>
<tr>
<td>Staphylococcus phage P954</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 1</td>
</tr>
<tr>
<td>Staphylococcus phage PH15</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 2</td>
</tr>
<tr>
<td>Staphylococcus phy2958PVL</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 3</td>
</tr>
<tr>
<td>Staphylococcus phage phiETA2</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 4</td>
</tr>
<tr>
<td>Staphylococcus phage phiETA3</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 5</td>
</tr>
<tr>
<td>Staphylococcus phage phiETA</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 6</td>
</tr>
<tr>
<td>Staphylococcus phage phiMR11</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 7</td>
</tr>
<tr>
<td>Staphylococcus phage phiMR25</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 8</td>
</tr>
<tr>
<td>Staphylococcus phage phiNM1</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane bacilliform IM virus</td>
</tr>
<tr>
<td>Staphylococcus phage phiNM3</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane bacilliform Mor virus</td>
</tr>
<tr>
<td>Staphylococcus phage phiP68</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane bacilliform virus</td>
</tr>
<tr>
<td>Staphylococcus phage phiPVL108</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane streak Egypt virus-Giza</td>
</tr>
<tr>
<td>Staphylococcus phage phiPVL-CN125</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane streak Reunion virus</td>
</tr>
<tr>
<td>Staphylococcus phage phiSauS-IPLA35</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sugarcane streak virus-Natal</td>
</tr>
<tr>
<td>Staphylococcus phage phiSauS-IPLA88</td>
<td>Streptococcus prophage phiSLT</td>
<td>Suid herpesvirus 1</td>
</tr>
<tr>
<td>Staphylococcus phage phiSLT</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus islandicus filamentous virus</td>
</tr>
<tr>
<td>Staphylococcus phage PT1028</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus islandicus rod-shaped virus 1</td>
</tr>
<tr>
<td>Staphylococcus phage ROSA</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus islandicus rod-shaped virus 2</td>
</tr>
<tr>
<td>Staphylococcus phage SAP-26</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus spindle-shaped virus 4</td>
</tr>
<tr>
<td>Staphylococcus phage SAP2</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus spindle-shaped virus 5</td>
</tr>
<tr>
<td>Staphylococcus phage Twort</td>
<td>Streptococcus prophage phiSLT</td>
<td>Sulfolobus spindle-shaped virus 5</td>
</tr>
<tr>
<td>Staphylococcus phage X2</td>
<td>Streptococcus prophage phiSLT</td>
<td>Subterranean clover stunt virus DNA 1</td>
</tr>
</tbody>
</table>

32
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
- Sulfolobus 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 Bengal 2008 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 Banglore virus-Ban5 satellite DNA beta
- Tomato leaf curl Banglore-virus
- Tomato leaf curl Bangladesh virus
- Tomato leaf curl Cameroon virus-CameroonBueaOkra2008
- Tomato leaf curl Cebu virus DNA-A
- 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-Taiwan Hsinchu 2005 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-Nigeria 2006
- 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-curly top virus
Tomato rugose mosaic virus DNA A
Tomato rugose mosaic virus DNA B
Tomato severe leaf curl virus DNA A
Tomato severe leaf curl virus DNA B
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 douroucouli 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
Trichodesmatis spinulosas-associated polyomavirus
Trichoplusia ni ascovirus 2c
Trichoplusia ni SNPV
Tupaiid herpesvirus 1
Turkey adenovirus A
Turnip curly top virus
TYLCCNV-Y322 satellite DNA beta
Urochloa streak virus
Vaccinia virus
Variola virus
The list of viral and bacterial strains includes:

- Velvet bean severe mosaic virus DNA A
- Velvet bean severe mosaic virus DNA B
- Vernonia yellow vein betasatellite
- Vernonia yellow vein virus DNA-A
- Vibrio phage fs1
- Vibrio phage fs2
- Vibrio phage K139
- Vibrio phage kappa
- Vibrio phage KSF-1phi
- Vibrio phage KVP40
- Vibrio phage N4
- Vibrio phage VEJphi
- Vibrio phage Vf12
- Vibrio phage VfO3K6
- Vibrio phage VfO4K68
- Vibrio phage VGJphi
- Vibrio phage VHML
- Vibrio phage VP2
- Vibriophage VP4
- Vibrio phage VP5
- Vibrio phage VP882
- Vibrio phage VP93
- Vibrio phage VpV262
- Vibrio phage VSK
- Watermelon chlorotic stunt virus DNA A
- Watermelon chlorotic stunt virus DNA B
- Wheat dwarf virus
- Wisadula golden mosaic St Thomas Virus DNA A
- Wisadula golden mosaic St Thomas Virus DNA B
- Woodchuck hepatitis virus
- Xanthomonas phage phiL7
- Xanthomonas phage Xop411
- Xanthomonas phage Xp10
- Xanthomonas phage Xp15
- Xenopus laevis endogenous retrovirus Xen1
- Xestia c-nigrum granulovirus
- Xylella phage Xfas53
- Yaba-like disease virus
- Yaba monkey tumor virus
- Yersinia pestis phage phiAl122
- Yersinia phage Berlin
- Yersinia phage L-413C
- Yersinia phage phiYeO3-12
- Yersinia phage PY54
- Zinnia leaf curl disease associated sequence virion
- Zinnia leaf curl virus-associated DNA beta
### Appendix 3: Bacterial Database

| Acaryochloris marina MBIC11017 | Actinosynnema mirum DSM 43827 |
| Acetobacter pasteurianus IFO 3283 01 | Aerococcus urinae ACS 120 V Col10a |
| Acetohalobium arabaticum DSM 5501 | Aeromonas hydrophila subsp hydrophila ATCC 7966 |
| Acholeplasma laidlawii PG 8A | Aeromonas salmonicida subsp salmonicida A449 |
| Achromobacter xylosoxidans A8 | Aeropyrum pernix K1 |
| Acidaminococcus fermentans DSM 20731 | Aggregatibacter actinomycetemcomitans D115 1 |
| Acidaminococcus intestini RyC MR95 | Aggregatibacter aphrophilus NJ8700 |
| Acidianus hospitalis W1 | Agrobacterium fabrum subsp C58 |
| Acidilobus saccharovorans 345 15 | Agrobacterium radiobacter K84 |
| Acidiphilium cryptum JF 5 | Agrobacterium sp H13 3 |
| Acidiphilium multivorum AIU301 | Agrobacterium vitis S4 |
| Acidithiobacillus caldus SM 1 | Akkermansia muciniphila ATCC BAA 835 |
| Acidithiobacillus ferrivorans SS3 | Alcanivorax borkumensis SK2 |
| Acidithiobacillus ferrooxidans ATCC 23270 | Alicyclobacillus acidocaldarius subsp acidocaldarius DSM 446 |
| Acidobacterium capsulatum ATCC 51196 | Aliivibrio salmonicida LFI1238 |
| Acidothermus cellulolyticus 11B | Alkalilimnicola ehrlichii MLHE 1 |
| Acidovorax avenae subsp avenae ATCC 19860 | Alkaliphilus metalliredigens QYM5 |
| Acidovorax citrulli AAC00 1 | Alkaliphilus oremlandii OHILAs |
| Acidovorax ebreus TPSY | Allochromatium vinosum DSM 180 |
| Acidovorax sp JS42 | Alteromonas macleodii str Deep ecotype |
| Aciduliprofundum boonei T469 | Alteromonas sp SN2 |
| Acinetobacter baumannii 1656 2 | Aminobacterium colombiense DSM 12261 |
| Acinetobacter calcoaceticus PHEA 2 | Ammonifex degensii KC4 |
| Acinetobacter oleivorans DR1 | Amycolatopsis mediterranei U32 |
| Acinetobacter sp ADP1 | Amycolicicoccus subflavus DQS3 9A1 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Anaerabacter variabilis ATCC 29413 |
| Actinobacillus succinogenes 130Z | Anaerococcus prevotii DSM 20548 |
| Actinoplanes sp SE50 110 | Anaerolinea thermophila UNI 1 |
| Actinophrye acidivorans DSM 10331 | Anaeromyxobacter dehalogenans 2CP 1 |
| Acidiphilium cryptum JF 5 | Anaeromyxobacter sp Fw109 5 |
| Acidiphilium multivorum AIU301 | Anaplasma centrale str Israel |
| Acidithiobacillus caldus SM 1 | Anaplasma marginale str Florida |
| Acidithiobacillus ferrivorans SS3 | Anaplasma phagocytophilum HZ |
| Acidithiobacillus ferrooxidans ATCC 23270 | Anoxybacillus flavithermus WK1 |
| Acidobacterium capsulatum ATCC 51196 | Aquifex aeolicus VF5 |
| Acidothermus cellulolyticus 11B | Arcanobacterium haemolyticum DSM 20595 |
| Acidovorax avenae subsp avenae ATCC 19860 | Archaeoglobus fulgidus DSM 4304 |
| Acidovorax citrulli AAC00 1 | Archaeoglobus profundus DSM 5631 |
| Acidovorax ebreus TPSY | Archaeoglobus veneficus SNP6 |
| Acidovorax sp JS42 | Arcobacter butzleri RM4018 |
| Aciduliprofundum boonei T469 | Arcobacter nitrofigilis DSM 7299 |
| Acinetobacter baumannii 1656 2 | Arcobacter sp L |
| Acinetobacter calcoaceticus PHEA 2 | Aromatoleum aromaticum EbN1 |
| Acinetobacter oleivorans DR1 | Arthrobacter arilaitensis Re117 |
| Acinetobacter sp ADP1 | Arthrobacter aurescens TC1 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Arthrobacter chlorophenolicus A6 |
| Actinobacillus succinogenes 130Z | Arthrobacter phenanthrenivorans Sphe3 |
| Actinoplanes sp SE50 110 | Arthrobacter sp FB24 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Aster yellows witches broom phytoplasma AYWB |
| Actinobacillus succinogenes 130Z | Asticcacaulis excentricus CB 48 |
| Actinoplanes sp SE50 110 | Atopobium parvulum DSM 20469 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Azohorizobium caulindans ORS 571 |
| Actinobacillus succinogenes 130Z | Azospirillum sp B510 |
| Actinobacillus succinogenes 130Z | Azotobacter vinelandii DJ |
| Actinoplanes sp SE50 110 | Bacillus amyloliquefaciens DSM 7 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Bacillus antracis str Ames |
SIANN: Strain Identification by Alignment to Near Neighbors

Bacillus atrophaeus 1942
Bacillus cellulosilyticus DSM 2522
Bacillus cereus 03BB102
Bacillus claussii KSM K16
Bacillus coagulans 26
Bacillus cytotoxicus NVH 98
Bacillus halodurans C125
Bacillus licheniformis DSM 13
ATCC 14580
Bacillus megaterium DSM 319
Bacillus mycoides DSM 2048
Bacillus pseudofirmus OF4
Bacillus pumilus SAFR 032
Bacillus subtilis BSn5
Bacillus thuringiensis serovar berliner ATCC 10792
Bacillus weihenstephanensis KBA4
Bacteriovorax marinus SJ
Bacteroides fragilis YCH46
Bacteroides helcogenes P 36
Bacteroides salanitronis DSM 18170
Bacteroides thetaiotaomicron VPI 5482
Bacteroides vulgatus ATCC 8482
Bartonella bacilliformis KC583
Bartonella clarridgeiae 73
Bartonella grahamii as4aup
Bartonella henselae str Houston 1
Bartonella quintana str Toulouse
Bartonella tribocorum CIP 105476
Baumannia cicadellinicola str Hc Homalodisca coagulata
Bdellovibrio bacteriovorus HD100
Beijerinckia indica subsp indica ATCC 9039
Beutenbergia cavernae DSM 12333
Bifidobacterium adolescentis ATCC 15703
Bifidobacterium animalis subsp animals ATCC 25527
Bifidobacterium bifidum S17
Bifidobacterium breve ACS 071 V Sch8b
Bifidobacterium dentium Bd1
Bifidobacterium longum DJO10A
Blattabacterium sp Blattella germanica str Bge
Bordetella avium 197N
Bordetella bronchiseptica RB50
Bordetella parapertussis 12822
Bordetella pertussis CS
Bordetella petrii DSM 12804
Borrelia afzelii PKo
Borrelia bissetti DN127
Borrelia burgdorferi B31
Borrelia duttonii Ly
Borrelia garinii PBl
Borrelia hermsii DAH
Borrelia recurrentis A1
Borrelia turicatae 91E135
Brachybacterium faecium DSM 4810
Brachyspira hydysenteriae WA1
Brachyspira intermedia PWS A
Brachyspira murdochii DSM 12563
Brachyspira pilosicoli 95 1000
Bradyrhizobium japonicum USDA 110
Bradyrhizobium sp BTAi1
Brevibacillus brevis NBRC 100599
Brevundimonas subvibrioides ATCC 15264
Brucella abortus A13334
Brucella canis ATCC 23365
Brucella melitensis bv 1 str 16M
Brucella microti CCM 4915
Brucella ovis ATCC 25840
Brucella pinnipedialis B2 94
Brucella suis 1330
Buchnera aphidicola str APS
Acyrthosiphon pisum
Burkholderia ambifaria AMMD
Burkholderia cenocepacia HI2424
Burkholderia gladioli BSR3
Burkholderia glumae BGR1
Burkholderia mallei ATCC 23344
Burkholderia multivorans ATCC 17616
Burkholderia phymatum STM815
Burkholderia phytofirmans PsJN
Burkholderia pseudomallei 668
Burkholderia rhizoxinica HKI 454
Burkholderia sp 383
Burkholderia thailandensis E264
Burkholderia vietnamiensis G4
Burkholderia xenovorans LB400
Butyrivibrio proteoclasticus B316
Caldicellulosiruptor bescii DSM 6725
Caldicellulosiruptor hydrothermalis 108
Caldicellulosiruptor kstejanssioni 177R1B
Caldicellulosiruptor kronoskyensis 2002
Caldicellulosiruptor lactoaceticus 6A
Caldicellulosiruptor obsidiensis OB47
Caldicellulosiruptor owensensis OL
Caldicellulosiruptor saccharolyticus DSM 8903
Caldivirga maquilingensis IC167
Campylobacter concisus IC13826
Campylobacter curvus 52592
Campylobacter fetus subsp fetus 82 40
Campylobacter fetus subsp
Campylobacter jejuni subsp jejuni 81116
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 Methylomirabilis oxyfera
Candidatus Midichloria mitochondrii IricVA
Candidatus Moranella endobia PCIT
Candidatus Nitrosipira delphii
Candidatus Pelagibacter sp IMCC9063
Candidatus Phytoplasma australiensis
Candidatus Protochlamydia amoebophila 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 Vescomyosocius okutanii HA
Candidatus Zinderia insecticola CARI
Carnobacterium cretonus CB15
Carnobacterium segnis ATCC 21756
Carnobacterium sp K31
Cellulomonas fimic ATCC 484
Cellulomonas flavigena DSM 20109
Cellulophaga algicola DSM 14237
Cellulophaga lytica DSM 7489
Cellvibrio gibris ATCC 13127
Cellvibrio japonicus Ueda107
Cenarchaeum symbiosum A
Chelatiorans sp BNC1
Chitinophaga pinensis DSM 2588
Chlamydia muriarum Nigg
Chlamydia trachomatis 434 Bu
Chlamydophila abortus S26 3
Chlamydophila caviae G01
Chlamydophila felis F 56
Chlamydophila pecorum E58
Chlamydophila pneumoniae CWL029
Chlamydophila psittaci 6BC
Chlorobaculum parvum NCIB 8327
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 thalassium 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 NCIMB 8052
Clostridium botulinum A str ATCC 3502
Clostridium cellulolyticum H10
Clostridium cellulovorans 743B
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
Comamonas testosteroni CNB 2
Conexibacter woesei DSM 14684
Caprothermobacter proteolyticus DSM 5265
Coraliomargarita 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 z3032
Cryptobacterium curtum DSM 15641
Cupriavidus metallidurans CH34
Cupriavidus necator N 1
Cupriavidus taiwanensis LMG 19424
cyanobacterium UCYN A
Cyanotohece sp ATCC 51142
Cyclobacterium marinum 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
Delftia acidovorans SPH 1
Delftia sp Cst 4
Denitrovibrio acetiphilus DSM 12809
Desulfbacter baarsii DSM 2075
Desulfatabiculum alkenivorans AK 01
Desulfobacterium hafniense YS1
Desulforbacinca acetoxidans DSM 11109
Desulfofobacterium autotrophicum HRM2
Desulfobulbus propionicus DSM 2032
Desulfococcus oleovorans Hxd3
Desulfotahialobium retbaense DSM 5692
Desulfomicrobium baculatum DSM 4028
Desulfotalea psychrophila LSv54
Desulfotomaculum acetoxidans DSM 771
Desulfotomaculum carboxydovorans CO 1 SRB
Desulfotomaculum kuznetsovi DSM 6115
Desulfotomaculum reducens MI 1
Desulfotomaculum ruminis DSM 2154
Desulfuromonas 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
Delftia acidovorans SPH 1
Delftia sp Cst 4
Denitrovibrio acetiphilus DSM 12809
Desulfbacter baarsii DSM 2075
Desulfatabiculum alkenivorans AK 01
Desulfobacterium hafniense YS1
Desulforbacinca acetoxidans DSM 11109
Desulfofobacterium autotrophicum HRM2
Desulfobulbus propionicus DSM 2032
Desulfococcus oleovorans Hxd3
Desulfotahialobium retbaense DSM 5692
Desulfomicrobium baculatum DSM 4028
Desulfotalea psychrophila LSv54
Desulfotomaculum acetoxidans DSM 771
Desulfotomaculum carboxydovorans CO 1 SRB
Desulfotomaculum kuznetsovi DSM 6115
Desulfotomaculum reducens MI 1
Desulfotomaculum ruminis DSM 2154
Desulfuromonas 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
Delftia acidovorans SPH 1
Delftia sp Cst 4
Denitrovibrio acetiphilus DSM 12809
Desulfbacter baarsii DSM 2075
Desulfatabiculum alkenivorans AK 01
Desulfobacterium hafniense YS1
Desulforbacinca acetoxidans DSM 11109
Desulfofobacterium autotrophicum HRM2
Desulfobulbus propionicus DSM 2032
Desulfococcus oleovorans Hxd3
Desulfotahialobium retbaense DSM 5692
Desulfomicrobium baculatum DSM 4028
Desulfotalea psychrophila LSv54
Desulfotomaculum acetoxidans DSM 771
Desulfotomaculum carboxydovorans CO 1 SRB
Desulfotomaculum kuznetsovi DSM 6115
Desulfotomaculum reducens MI 1
Desulfotomaculum ruminis DSM 2154
SIANN: Strain Identification by Alignment to Near Neighbors

Erwinia amylovora ATCC 49946
Erwinia billingiae Eb661
Erwinia pyrifoliae DSM 12163
Erwinia sp Ejp617
Erwinia tasmaniensis Et1 99
Erysipelothrix 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 RT17 B1
Fibrobacter succinogenes subsp succinogenes 585
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
Gallibacterium anatis UMN179
Gallionella capsiferriformans ES 2
gamma proteobacterium Hdn1
Gardnerella vaginalis 409 05
Gemmatimonas aurantiaca T 27
Geobacillus kaustophilus HTA426
Geobacillus sp C56 T3
Geobacillus thermodenitrificans NG80 2
Geobacillus thermoglucosidasi 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 RI4
Geodermatophilus obscurus DSM 43160
Glaciecola nitratireducens FR1064
Glaciecola sp 4H 3 7YE 5
Gloeobacter 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 tundrica
Haemophilus ducreyi 35000HP
Haemophilus influenzae 10810
Haemophilus parainfluenzae T311
Haemophilus parasuis SH0165
Haemophilus somnus 129PT
Hahella chejuensis KCTC 2396
Halalkalicoccus jeotgali B3
Halanaerobium hydrogeniformans
Haliangium ochraceum DSM 14365
Halicosmenobacter hydrossis DSM 1100
Halocarcula hispanica ATCC 33960
Halocarcula marismortui ATCC 43049
Halobacterium sp NRC 1
Halofex volcanii DS2
Halogeometricum borinquense DSM 11551
Halomonas elongata DSM 2581
Halophilic archaeon DL31
Halopiger xanaduensis SH 6
Halorhodospira walsbyi C23
Halorhabdus utahensis DSM 12940
Halorhodospira halophila SL1
Halorbrum lacsusprofundi ATCC 49239
Haloterrigena turkmenica DSM 5511
Halothermothrix orenii H 168
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

<table>
<thead>
<tr>
<th>Species/Strain</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helicobacter modesticaldum Ice1</td>
<td>DSM 9663</td>
</tr>
<tr>
<td>Herbaspirillum seropedicae SmR1</td>
<td>DSM 18828</td>
</tr>
<tr>
<td>Herminimonas arsenicoxydans</td>
<td>DSM 18837</td>
</tr>
<tr>
<td>Herpetosiphon aurantiacus</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hydrogenobacter thermophilus TK 6</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hydrogenobaculum sp YO4AA5</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hyperthermus butylicus DSM 5466</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hyphomicrobium denitrificans ATCC 51888</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hyphomicrobium sp</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Hyphomonas neptunium ATCC 15444</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Idiomarina loihiensis L2TR</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Ignicoccus hospitalis KIN4 I</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Ignisphaera aggregans DSM 17230</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Ilyobacter polytropus DSM 2926</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Intrasporangium calvum DSM 43043</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Isopetercola variabilis 225</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Isosphaera pallida ATCC 43644</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Jannaschia sp CCS1</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Janthinobacterium sp Marseille</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Jonesia denitrificans DSM 20603</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kangiella koreensis DSM 16069</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Ketogulonicigenium vulgare WSH 001</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kineococcus radiotolerans SRS3021</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kitasatospora setae KM 6054</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Klebsiella oxytoca KCTC 1686</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Klebsiella pneumoiae 342</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Klebsiella variicola At 22</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kocuria rhizophila DC2201</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kosmotoga olearia TBF 1951</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kribbelia flavida DSM 17836</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Krokinobacter sp 4H 3 7 5</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kyrtococcus tusciae DSM 2912</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Kytococcus sedentarius DSM 20547</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lacinutrix sp 5H 3 7 4</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus acidophilus NCFM</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus amylovorus GRL 1112</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus brevis ATCC 367</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus buchneri NRRL B 30929</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus casei ATCC 334</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus delbrueckii subsp bulgaricus ATCC 11842</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus fermentum IFO 3956</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus gasseri ATCC 33323</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus helveticus DPC 4571</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus johnsonii NCC 533</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus kefiranofaciens ZW3</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus plantarum subsp plantarum ST III</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus reuteri DSM 20016</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus rhamnosus ATCC 8530</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus ruminis ATCC 27782</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus sakei subsp sakei 23K</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus salivarius UCC118</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactobacillus sanfranciscensis TMW 11304</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactococcus garvieae ATCC 49156</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lactococcus lactis subsp cremoris NZ9000</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Laribacter hongkongensis HLHK9</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lawsonia intracellularis PHE MNI 00</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leadbetterella byssophila DSM 17132</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Legionella longbeachae NSW150</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Legionella pneumophila subsp pneumophila ATCC 43290</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leifsonia xyi subsp xyi str CTC807</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leptospira biflexa serovar Patoc strain Patoc 1 Ames</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leptospira borgpetersenii serovar Hardjo bovis str L550</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leptospira interrogans serovar Copenhageni str Fiocruz L1 130</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leptothrix cholodnii SP 6</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leptotrichia buccalis C 1013 b</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leuconostoc citreum KM20</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leuconostoc gasicomitatum LMG 18811</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leuconostoc mesenteroides subsp mesenteroides ATCC 8293</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Leuconostoc sp C2</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Listeria innocua Clip11262</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Listeria ivanovii</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Listeria monocytogenes EGD</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Listeria seeligeri serovar 1 2b str SLCC3954</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Listeria welshimeri serovar 6b str SLCC5334</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Lysinibacillus sphaericus C3 41</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Macroccocus caseolyticus JCS5402</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Magnetococcus marinus MC 1</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Magnetospirillum magneticum AMB 1</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Mahella australiensis 50 1 BON</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Mannheimia succiniciproducens MBEL5SE</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Maribacter sp HTCC2170</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Maricaulis maris MCS10</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Marinibacter adhaerens HP15</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Marinobacter aquaeolei VT8</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Marinomonas mediterranea MMB 1</td>
<td>DSM 19830</td>
</tr>
<tr>
<td>Organism Name</td>
<td>Strain Information</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Marinomonas posidonica</td>
<td>Po 181</td>
</tr>
<tr>
<td>Marinomonas sp</td>
<td>MWYL1</td>
</tr>
<tr>
<td>Marivirga tractuosa</td>
<td>DSM 4126</td>
</tr>
<tr>
<td>Melotheimurus ruber</td>
<td>DSM 1279</td>
</tr>
<tr>
<td>Melotheimurus silvanus</td>
<td>DSM 9946</td>
</tr>
<tr>
<td>Melliscoccus plutoinus</td>
<td>ATCC 35311</td>
</tr>
<tr>
<td>Mesosphaera flororum</td>
<td>L1</td>
</tr>
<tr>
<td>Mesolithicum ciceri</td>
<td>biovar biserrulae WSM1271</td>
</tr>
<tr>
<td>Mesolithicum loti</td>
<td>MAFF303099</td>
</tr>
<tr>
<td>Mesolithicum opportunistum</td>
<td>WSM2075</td>
</tr>
<tr>
<td>Metallosphaera cuprina</td>
<td>Ar 4</td>
</tr>
<tr>
<td>Metallosphaera sedula</td>
<td>DSM 5348</td>
</tr>
<tr>
<td>Methanobacterium sp</td>
<td>AL 21</td>
</tr>
<tr>
<td>Methanobrevibacter ruminantium</td>
<td>M1</td>
</tr>
<tr>
<td>Methanocaldococcus fervens</td>
<td>AG86</td>
</tr>
<tr>
<td>Methanocaldococcus infernus</td>
<td>ME</td>
</tr>
<tr>
<td>Methanocaldococcus jannaschii</td>
<td>DSM 2661</td>
</tr>
<tr>
<td>Methanocaldococcus sp</td>
<td>FS406 22</td>
</tr>
<tr>
<td>Methanocaldococcus vulcanius</td>
<td>M7</td>
</tr>
<tr>
<td>Methanocella arvoryzae</td>
<td>MRE50</td>
</tr>
<tr>
<td>Methanocella paludicola</td>
<td>SANAE</td>
</tr>
<tr>
<td>Methanococcoides burtonii</td>
<td>DSM 6242</td>
</tr>
<tr>
<td>Methanococcus aeolicus</td>
<td>Nankai 3</td>
</tr>
<tr>
<td>Methanococcus maripaludis</td>
<td>S2</td>
</tr>
<tr>
<td>Methanococcus vanielli</td>
<td>SB</td>
</tr>
<tr>
<td>Methanococcus voltae</td>
<td>A3</td>
</tr>
<tr>
<td>Methanocorpusculum labeirnum</td>
<td>Z</td>
</tr>
<tr>
<td>Methanoculleus marini</td>
<td>JR1</td>
</tr>
<tr>
<td>Methanohalobium evestigatum</td>
<td>Z 7303</td>
</tr>
<tr>
<td>Methanohalophilus mahii</td>
<td>DSM 5219</td>
</tr>
<tr>
<td>Methanoplanus petrolearius</td>
<td>DSM 11571</td>
</tr>
<tr>
<td>Methanopyrus kandleri</td>
<td>AV19</td>
</tr>
<tr>
<td>Methanoregula boonei</td>
<td>6A8</td>
</tr>
<tr>
<td>Methanoseta concili</td>
<td>GP6</td>
</tr>
<tr>
<td>Methanoseta thermophila</td>
<td>PT</td>
</tr>
<tr>
<td>Methanosalum zhilinae</td>
<td>DSM 4017</td>
</tr>
<tr>
<td>Methanosarcina acetivorans</td>
<td>C2A</td>
</tr>
<tr>
<td>Methanosarcina barkeri</td>
<td>str Fusaro</td>
</tr>
<tr>
<td>Methanosarcina maezei</td>
<td>Go1</td>
</tr>
<tr>
<td>Methanosphaera stadmanae</td>
<td>DSM 3091</td>
</tr>
<tr>
<td>Methanosphaera palustris</td>
<td>E1 9c</td>
</tr>
<tr>
<td>Methanospirillum hungatei</td>
<td>JF 1</td>
</tr>
<tr>
<td>Methanothermobacter marburgensis</td>
<td>str Marburg</td>
</tr>
<tr>
<td>Methanothermobacter thermautotrophicus</td>
<td>str Delta H</td>
</tr>
<tr>
<td>Methanothermococcus okinawensis</td>
<td>IH1</td>
</tr>
<tr>
<td>Methanothermus fervidus</td>
<td>DSM 2088</td>
</tr>
<tr>
<td>Methanotriss igneus</td>
<td>Kol 5</td>
</tr>
<tr>
<td>Methylobacterium chloromethanicum</td>
<td>CM4</td>
</tr>
<tr>
<td>Methylobacterium extorquens</td>
<td>AM1</td>
</tr>
<tr>
<td>Methylo bacterium nodulans</td>
<td>ORS 2060</td>
</tr>
<tr>
<td>Methylobacterium populi</td>
<td>BJ001</td>
</tr>
<tr>
<td>Methylobacterium radiotolerans</td>
<td>JCM 2831</td>
</tr>
<tr>
<td>Methylococcus capsulatus</td>
<td>str Bath</td>
</tr>
<tr>
<td>Methylococcus capsulatus</td>
<td>str alcaliphilum</td>
</tr>
<tr>
<td>Methylococcus marinum</td>
<td>str methanica MC09</td>
</tr>
<tr>
<td>Methylococcus mobili</td>
<td>str JLLW8</td>
</tr>
<tr>
<td>Methylococcus versatilis</td>
<td>301</td>
</tr>
<tr>
<td>Methylovorus</td>
<td>glucosetrophic SIP3 4</td>
</tr>
<tr>
<td>Methylovorus</td>
<td>sp MP688</td>
</tr>
<tr>
<td>Micavibrio aeruginosavorus</td>
<td>ARL 13</td>
</tr>
<tr>
<td>Micromonospora aurantiaca</td>
<td>ATCC 27029</td>
</tr>
<tr>
<td>Micromonospora</td>
<td>sp L5</td>
</tr>
<tr>
<td>Mobiluncus curtisii</td>
<td>ATCC 43063</td>
</tr>
<tr>
<td>Moerella thermacetica</td>
<td>ATCC 39073</td>
</tr>
<tr>
<td>Moraxella catarrhalis</td>
<td>RH4</td>
</tr>
<tr>
<td>Muricauda rustringensis</td>
<td>DSM 13258</td>
</tr>
<tr>
<td>Mycobacterium abscessus</td>
<td>ATCC 19977</td>
</tr>
<tr>
<td>Mycobacterium africanum</td>
<td>GM041182</td>
</tr>
<tr>
<td>Mycobacterium avium</td>
<td>104</td>
</tr>
<tr>
<td>Mycobacterium bovis</td>
<td>AF2122 97</td>
</tr>
<tr>
<td>Mycobacterium canettii</td>
<td>CIPT 140010059</td>
</tr>
<tr>
<td>Mycobacterium gilvum</td>
<td>PYR GCK</td>
</tr>
<tr>
<td>Mycobacterium leprae</td>
<td>TN</td>
</tr>
<tr>
<td>Mycobacterium marinum</td>
<td>M</td>
</tr>
<tr>
<td>Mycobacterium rhodesiae</td>
<td>NBB3</td>
</tr>
<tr>
<td>Mycobacterium smegmatis</td>
<td>str MC2 155</td>
</tr>
<tr>
<td>Mycobacterium sp</td>
<td>JDM601</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
<td>H37Rv</td>
</tr>
<tr>
<td>Organism</td>
<td>Strain/Reference</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Mycobacterium ulcerans</td>
<td>Agy99</td>
</tr>
<tr>
<td>Mycobacterium vanbaalenii</td>
<td>PYR 1</td>
</tr>
<tr>
<td>Mycoplasma agalactiae PG2</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma arthritidis 158L3</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma bovis PG45</td>
<td></td>
</tr>
<tr>
<td>Mycoplasma capricolum subsp capricolum ATCC</td>
<td>27343</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</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 str 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 str 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>Nakamuraella 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>Natrialba 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</td>
<td></td>
</tr>
<tr>
<td>Neorickettsia sennetsu str Miyama</td>
<td></td>
</tr>
<tr>
<td>Niastella koreensis GR20 10</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 winogradskiy 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>Nitrosomonas eutropha C91</td>
<td></td>
</tr>
<tr>
<td>Nitrosomonas sp AL212</td>
<td></td>
</tr>
<tr>
<td>Nitrosopumilus 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>Nocardioctonoides sp JS614</td>
<td></td>
</tr>
<tr>
<td>Nocardiopsis dasonvillei subs 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 carboxidivorans OMA5</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</td>
<td></td>
</tr>
<tr>
<td>Owenweeksia hongkongensis</td>
<td></td>
</tr>
<tr>
<td>DSM 17368</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus mucilaginosus KNP414</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus polymyxia E681</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus sp JDR 2</td>
<td></td>
</tr>
<tr>
<td>Paenibacillus terrae 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>Parvulaculax bermudensis HTCC2503</td>
<td></td>
</tr>
<tr>
<td>Pasteurella multocida subsp multocida str Pm70</td>
<td></td>
</tr>
<tr>
<td>Pectobacterium atrosepticum SCR11043</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 propionicigenes DSM 2379</td>
<td></td>
</tr>
<tr>
<td>Pelodictyon phaeoclahtratiforme BU 1</td>
<td></td>
</tr>
<tr>
<td>Pelotomaculum thermopropionicigenes SI</td>
<td></td>
</tr>
<tr>
<td>Persephonella marina EX H1</td>
<td></td>
</tr>
<tr>
<td>Petrotopha mobilis SJ95</td>
<td></td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Phenylobacterium zucineum HLK1
Photobacterium profundum SS9
Photobacterium profundum subsp. asymbiotica ATCC 43949
Photobacterium luminescens subsp. laumondii TTO1
Picrophilus torridus DSM 6068
Planctomyces brasiliensis DSM 5305
Planctomyces limnophilus DSM 3776
Polaromonas naphthalenivorans CJ2
Polaromonas sp JS666
Polaromonas sp JS666
Polynucleobacter necessarius subsp. asymbioticus QLW P1DMWA 1
Polynucleobacter necessarius subsp. asymbioticus QLW P1DMWA 1
Porphyromonas asaccharolytica DSM 20707
Porphyromonas gingivalis W83
Prevotella dentiflora F0289
Prevotella melaninogenica ATCC 25845
Prevotella ruminicola 23
Pseudomonas aeruginosa subsp. brassicacearum NFM421
Pseudomonas entomophila L48
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 suwonesis 111
Psychrobacter arcticus 273.4
Psychrobacter cryohalolentis K5
Psychrobacter sp PRwf 1
Psychromonas ingrahamii 37
Puillimonas sp t7 7
Pyrococcus abyssi GE5
Pyrococcus furiosus DSM 3638
Pyrococcus horikoshii OT3
Pyrococcus sp NA2
Pyrococcus yayanosii CH1
Pyrolobus fumarii 1A
P. aerophilum IM2
Pseudomonas brasilicense CGA009
Pseudomonas sp AD4-1
P. aerophilum IM2
Pseudomonas aeruginosa PAO1
Propionibacterium acnes 6609
Propionibacterium freudenreichii subsp. shermanii CIRM BIA1
Proteus mirabilis HI4320
Propionibacterium atlantica T6c
Pseudoalteromonas haloplanktis TAC125
Pseudoalteromonas sp SM9913
Pseudogulbenkiania sp NH8B
Pseudomonas aeruginosa PAO1
Ramlibacter tataouinensis TTB310
Renibacterium salmoninarum ATCC 33209
Rhizobium etli CFN 42
Rhizobium leguminosarum bv. viciae 3841
Rhodobacter capsulatus SB 1003
Rhodococcus equi 103S
Rhodococcus erythropolis PR4
Rhodococcus jostii Rhodococcus jostii RHA1
Rhodococcus opacus B4
Rhodoferax ferrireducens T118
Rhodomicrobium vannielii ATCC 17100
Rhodopirellula baltica SH 1
Rhodopseudomonas palustris CGA009
Rhodospirillum centenum SW
Rhodospirillum rubrum ATCC 11170
Rhodothermus marinus DSM 4252
Rickettsia africae ESF 5
Rickettsia akari strain Harford
Rickettsia bellii RML 369 C
Rickettsia canadensis strain McKiel
Rickettsia conorii strain Malish 7
Rickettsia felis URRWXCa1
Rickettsia heliogiangensis 054
Rickettsia japonica YH
Rickettsia massiliæ MTU5
Rickettsia peacockii strain Rustic
Rickettsia prowazekii strain Madrid E
Rickettsia rickettsii strain Iowa
Rickettsia sibirica 246
Rickettsia slovaca 13B
Rickettsia typhi strain Wilmington
Riemerella anatipestifer ATCC 11845 DSM 15868
Robinitalea bifomata HTCC2501
Roseburia hominis A2 183
Roseiflexus castenholzii DSM 13941
Roseiflexus sp RS 1
<table>
<thead>
<tr>
<th>Strain Identity</th>
<th>Strain Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roseobacter denitrificans OCH 114</td>
<td>Shewanella piezotolerans WP3</td>
</tr>
<tr>
<td>Roseobacter litoralis Och 149</td>
<td>Shewanella putrefaciens CN 32</td>
</tr>
<tr>
<td>Rothia dentocariosa ATCC 17931</td>
<td>Shewanella sediminis HAW EB3</td>
</tr>
<tr>
<td>Rothia mucilaginosa DY 18</td>
<td>Shewanella sp ANA 3</td>
</tr>
<tr>
<td>Rubrobacter xylanophilus DSM 9941</td>
<td>Shewanella violacea DSS12</td>
</tr>
<tr>
<td>Ruegeria pomeroyi DSS 3</td>
<td>Shewanella woodyi ATCC 51908</td>
</tr>
<tr>
<td>Ruegeria sp TM1040</td>
<td>Shigella boydii Sb227</td>
</tr>
<tr>
<td>Ruminococcus albus 7</td>
<td>Shigella dysenteriae Sd197</td>
</tr>
<tr>
<td>Runella slithyformis DSM 19594</td>
<td>Shigella flexneri 2a str 301</td>
</tr>
<tr>
<td>Saccharomonospora viridis DSM 43017</td>
<td>Shigella sonnei Ss046</td>
</tr>
<tr>
<td>Saccharophagus degradans 2 40</td>
<td>Sideroxydans lithotrophicus ES 1</td>
</tr>
<tr>
<td>Saccharopolyspora erythraea NRRL 2338</td>
<td>Simkania negevensis Z</td>
</tr>
<tr>
<td>Salinibacter ruber DSM 13855</td>
<td>Sinorhizobium fredii NGR234</td>
</tr>
<tr>
<td>Salinispora arenicola CNS 205</td>
<td>Sinorhizobium medicae WSM419</td>
</tr>
<tr>
<td>Salinispora tropica CNB 440</td>
<td>Sinorhizobium melloti 1021</td>
</tr>
<tr>
<td>Salmonella bongori NCTC 12419</td>
<td>Slackia Heliotrinireducens DSM 20476</td>
</tr>
<tr>
<td>Salmonella enterica subsp enterica serovar Enteritidis str P125109</td>
<td>Sodalis glossinidius str morsitans</td>
</tr>
<tr>
<td>Sanguibacter keddiei DSM 10542</td>
<td>Sorangium cellulosum So ce56</td>
</tr>
<tr>
<td>Sebaldella termitidis ATCC 33386</td>
<td>Sphaerobacter thermophilus DSM 20745</td>
</tr>
<tr>
<td>Segniliparus rotundus DSM 44985</td>
<td>Sphaerochaeta coccoides DSM 17374</td>
</tr>
<tr>
<td>Selenomonas sputigena ATCC 35185</td>
<td>Sphaerochaeta globus str Buddy</td>
</tr>
<tr>
<td>Serratia plymuthica AS9</td>
<td>Sphaerocystis pleomorpha str Grapes</td>
</tr>
<tr>
<td>Serratia proteamaculans 568</td>
<td>Sphingobacterium sp 21</td>
</tr>
<tr>
<td>Serratia sp AS12</td>
<td>Sphingobium chlorophenolicum L 1</td>
</tr>
<tr>
<td>Serratia symbiotica str Cinara cedri</td>
<td>Sphingobium japonicum UT265</td>
</tr>
<tr>
<td>Shewanella amazonensis SB2B</td>
<td>Sphingobium sp SYK 6</td>
</tr>
<tr>
<td>Shewanella baltica OS155</td>
<td>Sphingomonas wittichii RW1</td>
</tr>
<tr>
<td>Shewanella denitrificans OS217</td>
<td>Sphingopyxis alaskensis RB2256</td>
</tr>
<tr>
<td>Shewanella frigidimarina NCIMB 400</td>
<td>Spirochaeta caldaria DSM 7334</td>
</tr>
<tr>
<td>Shewanella halifaxensis HAW EB4</td>
<td>Spirochaeta smaragdinae DSM 11293</td>
</tr>
<tr>
<td>Shewanella ISOthia PV 4</td>
<td>Spirochaeta thermophila DSM 6192</td>
</tr>
<tr>
<td>Shewanella oneidensis MR 1</td>
<td>Spirosoma linguale DSM 74</td>
</tr>
<tr>
<td>Shewanella pealeana ATCC 700345</td>
<td>Staphylothermus hellenicus DSM 12710</td>
</tr>
<tr>
<td>Stackebrandtia nassauensis DSM 44728</td>
<td>Staphylothermus marinus F1</td>
</tr>
<tr>
<td>Staphylococcus aureus subsp aureus JH1</td>
<td>Starkeya novella DSM 506</td>
</tr>
<tr>
<td>Staphylococcus camosus subsp camosus TM300</td>
<td>Stenotrophomonas maltophilia K279a</td>
</tr>
<tr>
<td>Staphylococcus epidermidis ATCC 12228</td>
<td>Stigmatella aurantiaca DW4 3 1</td>
</tr>
<tr>
<td>Staphylococcus gallolyticus subsp gallolyticus ATCC BAA 2069</td>
<td>Streptococcus agalactiae 2603V R</td>
</tr>
<tr>
<td>Streptococcus gordonii ser A</td>
<td>Streptococcus dysgalactiae subsp equisimilis ATCC 12394</td>
</tr>
<tr>
<td>Streptococcus equi subspecies zooepidemicus</td>
<td>Streptococcus gallolyticus subsp galloyticus ATCC BAA 2069</td>
</tr>
<tr>
<td>Streptococcus macedonicus ACA DC 198</td>
<td>Streptococcus gordoni II ser A</td>
</tr>
<tr>
<td>Streptococcus mitis B6</td>
<td>Streptococcus pseudopneumoniae ST556</td>
</tr>
<tr>
<td>Streptococcus mutans NN2025</td>
<td>Streptococcus oralis Uo5</td>
</tr>
<tr>
<td>Streptococcus parasanguinis ATCC 15912</td>
<td>Streptococcus parauberis KCTC 11537</td>
</tr>
<tr>
<td>Streptococcus pneumoniae ATCC 43144</td>
<td>Streptococcus pasteurianus ATCC 15912</td>
</tr>
</tbody>
</table>
SIANN: Strain Identification by Alignment to Near Neighbors

Streptococcus pyogenes M1 GAS
Streptococcus salivarius JIM8780
Streptococcus sanguinis SK36
Streptococcus suis 05ZYH33
Streptococcus thermophilus CNRZ1066
Streptococcus uberis 0140J
Streptomyces avermitilis MA4680
Streptomyces bingchenggensis 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
Sulfobacillus islandicus M1425
Sulfobacillus solfataricus P2
Sulfobacillus tokodaii str 7
Sulfuricurvum kujiense DSM 16994
Sulfurihydragenibium azorense Az Fu1
Sulfurihydragenibium sp YO3AOP1
Sulfurimonas autotrophica DSM 16294
Sulfurimonas denitrificans DSM 1251
Sulfurospirillum deleyianum DSM 6946
Sulfurovum sp NBC37 1
Syntrophobacter thermophilum IAM 14863
Synechocystis sp PCC 6803
Syntrophobacter fumaroxidans MPOB
Syntrophobutulus glycolicus DSM 8271
Syntrophomonas wolfeii subsp wolfei str Goettingen
Syntrophothermus lipocalidus DSM 12680
Syntrophus aciditrophicus SB Tannerella forsythia ATCC 43037
Taylorella equigenitalis MCE9
Tepidanaerobacter acetatoxidans Re1
Teredinibacter turnerae T7901
Terriglobus saanensis SP1PR4 Tetragenococcus halophilus
Thauera sp MZ1T
Theraerobacter marianensis DSM 12885
Theranaerovibrio acidaminovorans DSM 6589
Thermocella potens JR
Thermoanaerobacter brockii subsp finnii Ako 1
Thermoanaerobacter italicus Ab9
Thermoanaerobacterium thermosaccharolyticum DSM 571
Thermoanaerobacterium xylanolyticum LX 11
Thermoanaerobacter matthianii subsp matthianii str A3
Thermoanaerobacter pseudethanolicus ATCC 33223
Thermoanaerobacter sp X513
Thermoanaerobacter tengcongensis MB4
Thermoanaerobacter wiegelli R1881
Thermobaculum terrenum ATCC BAA 798
Thermobifida fusca YX
Thermobispora bisporea DSM 43833
Thermococcus barophilus MP
Thermococcus gammatolerans EJ3
Thermococcus kodakarensis KOD1
Thermococcus onnurineus NA1
Thermococcus sibiricus MM 739
Thermococcus sp 4557
Thermocrinis albus DSM 14484
Thermodesulfatator indicus DSM 15286
Thermodesulfobacterium sp OPB45
Thermodesulfobium narugense DSM 14796
Thermodesulfovibrio yellowstonii DSM 11347
Thermofilum pendens Hrk 5
Thermomicrobium roseum DSM 5159
Thermomonospora curvata DSM 43183
Thermoplasma acidophilum DSM 1728
Thermoplasma volcanium GSS1
Thermoproteus uzoniensis 768 20
Thermosediminibacter oceani 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 paraluciscuniculi 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 distributa DSM 14429
Vulcanisaeta moutnovskia 768 28
Waddlia chondrophila WSU 86 1044
Weeksella virosa DSM 16922
Weissella koreensis KACC 15510
Wigglesworthia glossinidiae endosymbiont of Glossina brevipalpis
Wolbachia endosymbiont of Culex quinquefasciatus Pel
Wolbachia sp wRi
Wolinella succinogenes DSM 1740
Xanthobacter autotrophicus Py2
Xanthomonas albilineans GPE PC73
Xanthomonas axonopodis pv citri str 306
Xanthomonas campestris pv campestris str 8004
Xanthomonas oryzae pv oryzae KACC 10331
Xenorhabdus bovenii 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
Zobellia galactanivorans
Zunongwangia profunda SM A87
Zymomonas mobilis subsp mobilis ATCC 10988