

Mauve Version History

Version 2.3.1 (2009-11-11)

New features:

- * Compatibility with older versions of Mac OS X

Bugs fixed:

- * SNP output now handles mixed case sequence files and gap characters
- * Aligner memory usage has been streamlined, resulting in reduced memory requirements for alignments with many genomes
- * Fix for bug that prevented reordering GenBank-format draft genomes on the command-line
- * Fixed a crash when generating mums
- * Report an error when sequences containing gaps are used as input
- * Remove extra newlines in the ortholog alignment output
- * Fixed an OpenJDK rectangle drawing bug

Version 2.3.0 (2009-06-16)

Major new features:

- * Reordering contigs of draft genomes, published as Rissman et al 2009
- * Export a list of Single Nucleotide Polymorphisms present in an alignment (thanks to Meg Woolfit for the suggestion)
- * Export a file listing the genome arrangement as a signed permutation
- * Export a list of positionally orthologous CDS, tRNAs, or misc_RNAs (thanks to Elizabeth Skippington for the suggestion)
- * Export a file containing alignments of orthologous CDS, including alignments of unannotated regions predicted to be orthologous (thanks to Sam Sheppard for the idea)

Other new features:

- * On Linux it is now possible to launch Mauve from outside the Mauve directory (thx Dongying Wu)
- * Option to disable drawing the mouse cursor highlighting orthologous regions so that figures for manuscripts can be rendered (thanks to Andrew Kropinski for the idea)

Bugs fixed:

- * An alignment would complete successfully but the viewer would fail to load it because the alignment contained invalid data in the backbone file. This bug was very commonly encountered in 2.2.0.
- * An alignment with progressiveMauve would crash with an error like "cga doesn't fit" or another error message, when run on multicore computers. This was a race condition in progressiveMauve's parallelism.

Known issues:

- * Parallelism has been disabled for all supported platforms to avoid a race condition bug.

Version 2.2.0 (2008-06-23)

New features:

- * New menu option to disable the red contig boundary lines
- * Progressive aligner sensitivity/accuracy improvements

Bugs fixed:

- * progressiveMauve was not sensitive enough to small rearrangements, although it did have high predictive value
- * progressiveMauve was not properly extending LCBs in the 2.1.x releases
- * progressiveMauve did not work on 64-bit linux (indefinite waiting message)
- * Versions 2.1.x did not render vector graphics when printing to PDF or EPS

- * seed families are too memory intensive and are now disabled by default
- * segmentation fault (crash) on linux/mac due to a bug in c++ sort functions
- * update check causes a stall during GUI launch when the network is inaccessible
- * building from source was difficult

Known issues:

- * Mauve 2.2.0 for Macintosh can not run in parallel on more than one CPU
- * Additional bugs can be found at http://sourceforge.net/tracker/?group_id=181544&atid=897627

Version 2.1.1 (2007-11-27)

Bugs fixed:

- * The alignment programs mauveAligner and progressiveMauve did not work on Mac due to a missing shared library (libgomp.dylib)
- * progressiveMauve would occasionally crash during alignment scoring and greedy breakpoint elimination due to a math bug
- * The sequence navigator did not properly highlight feature search results when launched from the menu. It did work when launched from the toolbar binoculars button.

Version 2.1.0 (2007-11-20)

New features:

- * A graphical interface to reorder and hide genomes from the display, and to change the reference genome
- * A new graphical interface to change the viewing style for genome comparisons
- * Parallelized alignment algorithms that run faster on multi-core CPUs like the Intel Core Duo and the AMD Opteron
- * 64-bit implementations for Windows and Mac OS X. 64-bit machines with enough RAM can align 20 or more bacterial genomes
- * Seed family searches. Several seed patterns can be used simultaneously to improve sensitivity on heavily diverged genomes

Bugs fixed:

- * Several progressiveMauve alignment crashes
- * A mauveAligner bug whereby the aligner never finished on some datasets
- * Many more. For a full list of known and fixed bugs please see http://sourceforge.net/tracker/?group_id=181544&atid=897627

Known issues:

- * Mauve 2.1.0 for Macintosh requires OS X version 10.4 or later
- * A bug exists which can crash Mauve on Linux and Mac OS X during the "recursive anchor search."
- * Mauve does not work with GNU java, which is the default java in some Linux distributions
- * Sun JRE 1.6.0 for Linux [has a bug](#) that breaks alignment loading (it's not my fault!). It may be necessary to manually open the alignment file using "File->Open" after making the alignment.

Version 2.0.0 (2007-01-26)

Mauve 2.0 includes an accurate new alignment algorithm called **Progressive Mauve**. The progressive alignment algorithm has been described in Aaron Darling's Ph.D. thesis, available from http://gel.ahabs.wisc.edu/~darling/darling_thesis.pdf

Progressive Mauve features:

- * **Substantially improved sensitivity on divergent genomes.** Progressive Mauve can "climb a phylogenetic ladder" such that a pair of genomes with > 1 substitution per site can be aligned if phylogenetic intermediate genomes are included in the alignment
- * **Segments that are present in subsets of the genomes will be aligned.** The original Mauve aligned only segments conserved in all genomes.
- * **A configurable substitution matrix dictates the LCB and alignment scores.** The substitution matrix gives Mauve additional power in distinguishing true homology from random sequence similarity. In the original algorithm LCBs were scored based on the sum of multi-MUM lengths
- * **A random-walk homology test identifies conserved "backbone" segments.** Random walk statistics similar to those used in Gapped-BLAST are used to identify pairwise conserved segments. Pairwise homologies are then combined by transitive homology relationships to identify segments conserved in two or more genomes.

- * **Progressive Mauve can scale to hundreds of genomes.** We have successfully aligned over a hundred poxvirus genomes, hundreds of Hepatitis C Virus genomes, and the genomes of over thirty enterobacteria.

Anna Rissman has contributed a new annotation search interface that tremendously improves browsing genomes with annotated sequence features. The new search interface can:

- * Jump directly to a particular numerical coordinate in a particular genome
- * Jump directly to an annotated feature of a particular name
- * Search for and list annotated features by any part of their name or qualifying value. Searchable fields include things like CDS name, product description, GO terms, database x-refs, and amino acid subsequences

Other new features:

- * The windows installer supports installations by users without administrative privileges and also supports multiple concurrent installations
- * The windows installer includes Java 6 for machines without an existing Java installation
- * Alignments from Progressive Mauve can be viewed in "backbone color" mode, which colorizes segments based on the combination of genomes in which they are conserved.

Bugs fixed:

- * Infinite looping during LCB extension should be fixed. LCB extension may still be slow for some datasets, with questionable benefit. It can be disabled in the Parameters tab of the "Align sequences.." window.
- * Several bugs were fixed that would cause sporadic crashes in MUSCLE 3.6.

Known bugs:

- * Bug tracking has moved to the sourceforge web site. Please use the following address to report bugs:
http://sourceforge.net/tracker/?group_id=181544&atid=897627

Version 1.3.0 (2006-05-24)

New features:

- * Rearrangement history analysis by generalized transposition/block interchange, code courtesy of Mike Tsai. Use the green/blue/black toolbar button to perform an analysis. For details on the inference algorithm see Yancopoulos, S., Attie, O and Friedberg, R. "Efficient sorting of genomic permutations by translocation, inversion and block interchange" *Bioinformatics* 2005 21(16):3340-3346
- * Rearrangement history through the GRIMM/MGR server now supported, use the red/yellow/black toolbar button.
- * No more limit on the number of input sequences. This feature comes at the cost of increased memory consumption. Just because the limit has been removed doesn't mean that the alignment algorithm will do a good job handling a large number of sequences.
- * The mauveAligner command-line tool has the new option --lcb-match-input to extract matches from an existing alignment and use them as anchors when computing a new alignment.
- * MUSCLE has been upgraded to version 3.6.
- * The Mac OS X version is now a Universal binary with support for both Intel and PowerPC macs.

Bugs fixed:

- * MUSCLE on Windows now runs faster and uses less memory.
- * On unix, Mauve will search \$PATH for mauveAligner.
- * Eliminated the obnoxious behavior where Mauve would repeatedly pop up the console window during the alignment process
- * Fixed bugs when viewing alignments after shifting the LCB minimum weight slider.

Version 1.2.3 (2005-10-20)

New features:

- * Display each component of multi-part features (e.g. a multiple exon gene) as a linked series of boxes
- * Display repeat_region annotations from GenBank files
- * Improved load time for previously viewed alignments using a disk-based alignment cache
- * Added a help menu item that will clear the on-disk alignment cache

- * Colorized the similarity plots, grey was starting to depress our entire lab.
- * Support for highlighting arbitrary regions of the sequence by holding the shift key while clicking and dragging on the display
- * Faster nucleotide display when zoomed in closely

Bug fixes:

- * Fixed GUI interface for setting the minimum seed size (Thanks to Bill Bruno and Todd Treangen)
- * Fixed mouseover, highlighting, and display alignment in the LCB display mode (used when the Full Alignment option is disabled)

Version 1.2.2 (2005-06-11)

New features:

- * Ability to set the reference genome
- * Displays gene name or locus tag for annotated features

Bug fixes:

- * Uses less memory for MUSCLE alignments by restricting the size of segments aligned by MUSCLE
- * Alignment editor shows original sequence coordinates when an aligned character is clicked
- * Handles GenBank files with a large number of sequence entries (e.g. C. hominis)

Version 1.2.1 (2005-05-05)

New features:

- * Experimental support for alignment editing using Cinema-MX Right click on an LCB to trigger the edit menu
- * Annotated features are shown in both forward and reverse complement orientation
- * When clicking an annotated feature, a database xref popup menu shows options for viewing the feature in known databases. Currently supported databases are Entrez Protein and ASAPdb. Please contact us about support for other databases.
- * Choice of either MUSCLE or Clustal in the GUI or on the command line with the --gapped-aligner switch

Bug fixes:

- * Database xref speed improvement
- * Features not shown when viewing large regions to improve speed
- * Fixed an off-by-one bug that crashed the viewer when loading Salmonella alignments
- * Eliminated extraneous MUSCLE warning messages on the Mac OS X console.

Known issues:

- * In some cases MUSCLE on Windows can consume lots of system resources. If you are experiencing problems we suggest using Clustal for the time being.

Version 1.2.0 (2005-04-20)

New features:

- * **Image export.** Mauve can export the current display view in JPEG, TIF, and PNG formats at a user-defined resolution.
- * **MUSCLE Alignment.** Mauve now calculates gapped alignments using the [MUSCLE](#) DNA alignment method. Previously Clustal-W was used to calculate gapped alignments. MUSCLE provides significantly improved speed and accuracy over Clustal-W
- * **Collinear genomes.** When it is known that the genomes being aligned have no rearrangements, select the Assume collinear genomes option in the "Align sequences" window to tell Mauve that it should find the single best LCB. On the command line, this behavior can be invoked with the --collinear switch.

Bug fixes:

- * Fixed a memory leak in LCB extension that could quickly consume all system memory.
- * The 1.1.0 Windows release used a large amount of memory because it was compiled for use with 160 sequences instead of the

usual 16.

- * Fixed a bug parsing GenBank sequences which had extra newlines in them

Known issues:

- * On Mac OS X the muscle aligner generates lots of warning messages about not being able to open /proc/meminfo. These messages can be ignored and will be eliminated in future releases.

Version 1.1.0 (2005-03-30)

New features:

- * Mauve can launch a web browser to show a web database entry for annotated features that have a /db_xref qualifier in their GenBank entry. Currently ASAPdb and NCBI's PubGene databases are supported. Other databases may be added upon request.
- * Mauve bundles a text console display for informational and error messages. It can be viewed by selecting Show console from the Help menu.
- * Alignment generation has been multithreaded so that alignments can be generated while another alignment is being viewed.
- * A Print Preview and Page Layout dialog boxes provide greater control over printing.

Bug fixes:

- * A bug in mauveAligner which resulting in a "bad alignment" error message has been fixed.
- * Printing has been fixed to provide accurate high resolution images. Output is vectorized so images printed to PDF documents (e.g. using Acrobat or PDF995) can be easily embedded in publications and have a high quality look.
- * Screen space is used more economically when viewing a large number of sequences, allowing at least six genomes to display on a 1024x768 screen. Previously only three genomes would fit on a screen of that size.
- * additional fixed bugs can be found in the Mantis tracker

Version 1.0.1 (2005-03-08)

Bug fixes:

- * Unable to align genomes on Mac OS X 10.3 (due to a missing library)
- * Unable to run Mauve under Linux (Mauve was not set as executable)
- * Unable to load genome annotations after constructing an alignment when the source GenBank files reside in a different directory than the alignment output
- * Lower memory profile. Mauve now uses less memory when reading alignments, making the large Drosophila alignments available on the web site viewable on computers with over 1GB RAM.

Version 1.0 (2005-03-01)

New features:

- * Display of annotated features such as CDS, rRNA, tRNA and misc_features from GenBank format files
- * Full support for Mac OS X
- * Display of chromosomal boundaries and contigs of incomplete genomes
- * Display of contig/chromosome name and sequence coordinates
- * Source code released under the GNU General Public License. Mauve is now free software and can be modified and redistributed freely (in accordance with the GPL).

2004-12-16

Bug fixes:

- * Inter-LCB islands were (still) incorrect and are now fixed. This bug was originally reported by Val Burland in Dec. 2003 but didn't get completely fixed until now.
- * Vertical only display resize fixed
- * Fixed a bug in the LCB info display when the LCB slider had been adjusted

New Features:

- * Sequence similarity browsing in the display
- * More sensitive alignments by using inexact seeds. This means Mauve can be used on more taxa simultaneously and also on more divergent taxa.
- * Display alignment in the similarity browser can be accomplished with a single mouse click

2004-07-08

Bug fixes:

- * Brought PDF documentation up-to-date

2004-06-30

Bug fixes:

- * Fixed several missing icon bugs in linux and windows releases
- * Mauve now requires Java 1.4 to function correctly, a warning message is displayed if Mauve detects an older version of Java

New features:

- * Automatic checking for new releases, and on Windows, downloading the new release and starting the installer
- * Releases now include a User Guide in PDF format and a link to the online guide

2004-05-18

Bug fixes:

- * Fixed a bug that would result in unreasonably large print jobs

New features:

- * Implemented interactive LCB weight adjustment. This feature allows dynamic adjustment of the LCB weight within Mauve using a slider control. As the minimum LCB weight is changed the display updates with the corresponding set of LCBs.
- * Added toolbar buttons for basic navigational controls such as scrolling and zooming the display. Also added a color mode selector.
- * Added keyboard shortcuts for open (ctrl+o), close (ctrl+w), print (ctrl+p), and quit (ctrl+q).

Important changes:

- * The up and down arrows no longer zoom in and out, instead use ctrl+up and ctrl+down

2004-04-26

Bug fixes:

- * Fixed a bug that allowed users to zoom in too far
- * Fixed a bug that kept the screen from clearing after selecting close from the File menu
- * Fixed a bug that would cause the display to become unaligned after using the context menu to align the display and subsequently zooming out.
- * Fixed a bug that raised an exception when no matching regions were found among the sequences being aligned
- * Changed the .txt file extension to default to FastA file format instead of raw sequence data
- * Fixed a subtle bug that would occasionally cause a "horrible error" to be reported

New features:

- * Improved the Align sequences dialog interface to give the user an easy set of default values with optional control over the alignment parameters. Also added an interface to set the backbone parameters.
- * Added the ability to change ordering of genomes in the display using drag and drop. To enter and leave reordering mode press the h key. When in reordering mode genomes can be dragged and dropped into the desired order.

2004-02-19

Bug fixes:

- * Fixed a bug in the island coordinate output that caused the end positions of inverted regions to be huge (wrong) numbers. Removed obsolete -d and --stats command line options. Added a --id-matrix or -t command line option to print out an identity matrix. Mauve now creates an identity matrix by default.

2003-12-09

Bug fixes:

- * Fixed another bug in the LCB extension code that occasionally resulted in incorrect alignments.

2003-12-06

Bug fixes:

- * Fixed a serious bug that caused the first sequence to be misaligned in regions with numerous gaps.

2003-11-04

Bug fixes:

- * Fixed a minor file format problem with XMFA (eXtended Multi-FastA) output. Fixed a bug with reverse complement backbone determination.

2003-10-17

Bug fixes:

- * Updated the GUI interface to match new command line options for the aligner. The GUI was behaving badly in the 2003-10-08 release. Full alignment is now selected by default, if no LCB weight is specified one will be chosen automatically.

2003-10-08

Bug fixes:

- * Final aligner debugging completed using thousands of test cases. Numerous bugs were squashed. The aligner should be (but isn't) rock solid now. Removed file I/O from the ClustalW code, resulting in vast speedups during recursive alignment. One more bug to kill.

2003-09-13

Bug fixes:

- * Debugged the LCB extension code, hopefully eliminating strange crashes from the previous release.

2003-09-08

New Features:

- * Added LCB extension code, allows Mauve to search for additional MUM anchors off the ends of LCBs. Increased the minimum anchor size to 5 b.p. to reduce spurious anchoring.

2003-08-15

New Features:

- * Added print menu item to the File menu - Clean-ups for Linux - Mauve for Linux is ready!

2003-08-13

Bug fixes:

- * Additional bugfixes for the recursive anchoring and alignment process. Will this software ever be bug free?

2003-08-02

Bug fixes:

- * Fixed some serious bugs in the production of gapped alignments from Mauve style MUM alignments. This bug affected island predictions and backbone predictions of previous releases.

New features:

- * Implemented Multi-FastA sequence input. If a single sequence file is given Mauve assumes it is a Multi-FastA file where each sequence entry corresponds to a sequence to align.

2003-07-31

New features:

- * Implemented backbone identification -- backbone is any region of sequence conserved among all genomes under study. Mauve's definition of backbone is any region longer than x base pairs that does not contain any gaps longer than x base pairs. x is specified as the island size in the "Align Sequences..." window.

2003-07-30

Bug fixes:

- * Fixed a segfault in SML binary search

New features:

- * Added end-gap penalties to clustal alignments

2003-07-28

Bug fixes:

- * Fixed a few problems with Clustal Alignment -- more to come

2003-07-25

Bug fixes:

- * Changed the default maximum Clustal alignment size to 9000 - Fixed numerous bugs related to recursive alignment and ClustalW alignment

New features:

- * Now outputs alignments in Extended Multi-FastA (xmfa) format

2003-06-12

Bug fixes:

- * Don't write a gapped alignment unless the user requests a Full alignment

2003-05-08

Bug fixes:

- * bugfixes for displaying large (> 2 G.B.) sequences

New features:

- * new ruler label style for easier reading

2003-05-05

Bug fixes:

- * Fixed DragNDrop for align sequences window

2003-04-28

Bug fixes:

- * Fixed Align sequences window disappearance

New features:

- * Added clustalW format alignment output (command line mauveAligner only)

2003-04-26

Bug fixes:

- * Fixed printing (broken in previous release)

New features:

- * Supports drag and drop of alignments to the alignment window and of sequence files to the sequences list

2003-04-21

Bug fixes:

- * Fix mouse click accuracy problem when zoomed in
- * don't clear alignment window after alignment

New features:

- * mauveAligner finds interLCB islands
- * copy the File chooser view to new alignment windows

2003-04-04

New features:

- * Have mauveAligner output islands, tree, and alignment
- * Implement high-resolution printing hack

2003-04-03

New features:

- * Add LCB color lines to link LCBs

2003-02-21

New features:

- * Initial Release - Packaged into an installer for windows