# MacVector 17.0.2

## for Mac OS X

## The online updater for this release is 143.5 MB in size

## You must be running MacVector 15.5.4 or later for this updater to work!

## **System Requirements**

MacVector 17.0 is supported on any Intel Macintosh running **Mac OS X 10.9** or higher. There are no other specific hardware requirements for MacVector – if your machine can run OS X 10.9 or above, it can run MacVector. A complete installation of MacVector 17.0 uses approximately 253 MB of disk space. Please note this release will NOT run on OS X 10.7 or earlier and is likely to have problems on OS X 10.8.

**ASSEMBLER NOTE**: If you are performing contig assembly using *MacVector with Assembler*, we recommend you have **at least** 2 GB of **FREE** RAM available on your machine. For any serious NGS work using phrap, velvet, SPAdes or bowtie, you should have at least 8 GB and preferably 16 GB or more for satisfactory performance.

### Installation and License Activation

You can choose to install MacVector in one of two ways; if you want to install MacVector for all users of the computer, simply drag the MacVector folder onto the Applications folder. You will be prompted for a system administrator account and password during this copy. If you don't have administrative privileges, or if you want to install it for just your own use, you can install MacVector in the /Applications/ folder in your own personal home directory. To do that, double-click on the

**OpenUserApplicationsFolder** icon to open the folder in a separate window, then drag the MacVector folder into that window.

Once installation is complete, when you first run MacVector you must enter a valid license owner, serial number and activation code, if one does not already exist on your computer. This information is usually sent by e-mail but is also printed on the inside of the CD sleeve if you opted to receive MacVector on physical media. If you previously installed an earlier version of MacVector and have a serial number with a maintenance end date of September 1<sup>st</sup> 2017 or later, MacVector 17 will automatically use your existing license and you will not be required to enter the details again. NOTE: one important change for MacVector 14.0 and later is that you can now optionally activate a license for your own use without requiring Administrative privileges.

## Changes for MacVector 17.0.2

#### **Bug Fixes**

Multi-segmented CDS features are now correctly translated in text output windows. Segmented alignments are now correctly displayed in the *Align to Reference* **Map** tab. The Base Composition text output now correctly calculates the "T" composition.

We have found a workaround for problems printing many graphical image views under macOS Mojave.

## Changes for MacVector 17.0.1

#### **Bug Fixes**

The *cDNA Alignment* algorithm in the Align to Reference function now generates results similar to MacVector 16.

A crash when hovering the mouse pointer over a tab bar with the <option> key held down has been fixed.

A problem opening saved Gibson Assembly projects on macOS versions prior to 10.14 has been resolved.

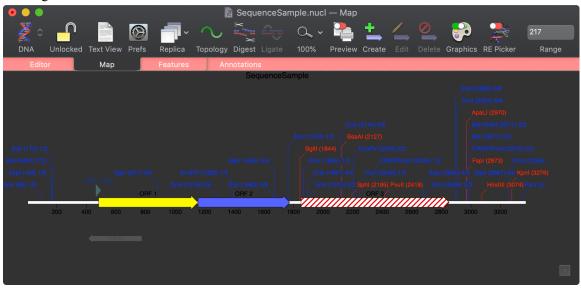
A variety of issues editing and selecting sequences in the *Multiple Sequence Alignment* Editor have been fixed.

When you select and ORF in the *Open Reading Frame* analysis results window, the corresponding region in the parental sequence is now selected.

## Changes for MacVector 17.0.0

#### Dark Mode Support

MacVector 17 has basic support for the macOS Mojave "Dark Mode". Many of the icons have been reworked and the background and foreground colors and the tab bars of most windows are now dynamic so that text and graphics are clear and legible under both Dark and Light modes.



#### **Restriction Enzyme Picker**

There is a new floating window that displays a list of Restriction Enzymes that cut the currently active sequence, even if those enzymes are not currently visible in the **Map** tab.

		R	estriction En:	zymes				
Filter	Filter			File				
Cuts: >= 0 and <= 8 Cuts: >=				Common Enzymes.renz 🗘				
Reset Filters   ? Select All   Deselect All Defaults   Save current set of Enzymes   Acc65i Acc65i-DCh   Acel Afel								
	Asel	AsiSI	att*1	att*2	att*3	att*4		
✓ attB1	✓ attB1.1	✓ attB2	✓ attB2.1	✓ attB3	✓ attB4	✓ attL1		
✓ attL2	✓ attL3	✓ attL4	✓ attP1	✓ attP2	✓ attP3	✓ attP4		
✓ attP5	✓ attR1	✓ attR2	✓ attR3	✓ attR4	Avrll	✓ BamHI		
Banl-DCM	Bfal	✓ BgIII	Blpl	Bmtl	BsaBI-DAM	BsaHI-DCM		
Bsml	BspDI	BspDI-DAM	BspEl	BspEI-DAM	BspHI	BspHI-DAM		
BsrGl	BstBl	BstZ17I	Bsu36l	Clal	Clal-DAM	CRISPR-righ		
./ Dral	FeePI	-/ FcoPV	C EDT	match these enzym		Hnal		

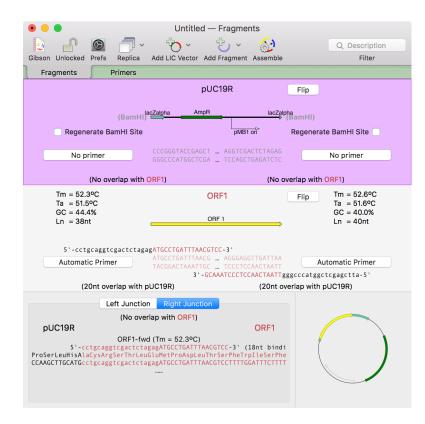
The entire interface is interactive, so you can move the sliders to filter by number of cuts or site size, or choose the end structure (5' overhang, 3' overhang, blunt) or toggle individual enzymes on and off. You can save any set of enzymes in a separate file for use in **Analyze** | **Restriction Enzymes...** analysis. You can show/hide the window when you are not interested in restriction enzymes using the **RE Picker** toolbar button (note there is also now a companion button to show/hide the **Graphics Palette**);



#### Gibson/Ligase Independent Cloning

There is a completely new interface and document type to handle a variety of ligaseindependent cloning strategies. The interface supports both the 5' exonuclease driven "Gibson" assembly technique as well as the T4 DNA Polymerase 3' exonuclease "Ligase Independent Cloning" approach. You can have MacVector automatically design primers for you, or you can provide custom primers you have designed yourself manually, or with popular internet tools such as the NEBuilder website. You can also just provide existing fragments with overlapping ends for MacVector to assemble. The interface also shows the exact structure of the junctions between fragments, including relevant CDS translations so that you can check you have protein fusions in the correct frame.

The interface is highly interactive – you can drag features into it from other sequence documents and even drag a restriction enzyme(s) into it from a vector to use the cut vector as a base for you cloning experiment. The recommended primers are listed in a separate tab that can be printed or copied to other applications, so you can document them before sending them off for synthesis. A simple click creates the new circular molecule once all ends have been accounted for. The entire project can be saved as a new document type, and any fragment can be separately opened in a normal sequence window either before or after applying PCR primers.



#### **Genome Comparison**

In recent years that has been an explosion of whole-genome sequencing projects. One common question coming out of this has been to ask "Exactly what are the genetic differences between my sequenced organism and another related strain?" MacVector to the rescue! A new **Analyze** | **Compare Genomes By Feature...** feature lets you see the differences between two annotated genomes in fine detail. The function takes every annotated feature from the source genome and looks for the presence of that feature in the comparison genome based on sequences are compared. The results are then tabulated to show identical, closely related, and weakly related features in separate tabs, with additional tabs for features that are completely missing and a "details" tab that shows the low-level alignment details for any matching pair of features. Hot-links in the result tabs let you quickly scroll the parent sequences to any individual feature of interest.

	🕺 Campylobacter jejuni 35925.nucl — Results	
Details		
/codon_s /db_xref /inferen /locus_t /product /protein /transl_	t(179918180028) art=1 GI:761233174 e=ab initio prediction:Prodigal:2.60 g=UC78_0175 hypothetical protein id=AP34688.1	
CDS compleme /codon_s /db_xref /locus_t /product /protein /transl_	G1:284925468 grCJSA_0160 putative lipoprotein id=A0c27820.1	
Aligned Length = 41 Identities = 36 (87.)		
Campylobacter jejuni Campylobacter jejuni	1 MKKIILFLLCVGFAFACSEHSHTDFKDLNKTEYNSQ 1 MKKIILFLLCVGFAFACSEHSHTDFKDLNKTEYNSQLKDKI	36 41
Raw alignment score	Unaligned residues = 14 213 (maximum raw score = 222) = 95.9% ting for unaligned residues) = 90.6%	
Campylobacter jejuni	179918 ATGAAAAAATAATTTATTTTACTATGCGTTGGCTTTGCTTTGCTTGC	79977
Campylobacter jejuni	172586 ATGAAAAAAAAAAATAATTTTAATTTTACTATGCGTTGGCTTTGCTTTGCTTGC	72645
Campylobacter jejuni		80037
Campylobacter jejuni		72705
Campylobacter jejuni	180038 18	80043

#### **Assembly Jobs**

You can now run multiple jobs in an assembly project and each job is encapsulated in its own job object in the Assembly Project window. This gives you much greater flexibility in organizing your assembly tasks and each job is given an intuitive name.

• • • Untitled 4 — Project								
- ta 🖞 🚰 🍇 👰 🗛 🙆 🗇 .	🖺 🖞 🤫 🖁	a 🗣 🚳	Q Name					
Add Reads Add Seqs Add Ref Add Contig Remove Reset Prefs Replica	Phred CrossMatch Phrap Bow	tie SPAdes Velvet	Filter					
Project Properties Coverage								
Name	Status Length	#	ClipR St Stop					
MannotatedGenome	REF 1618194	. 1	1618194					
Cj-RM35925_S23-1643_L001_R1_001.fastq	Ilumina Paired-e 301	854542						
Gj-RM35925_S23-1643_L001_R2_001.fastq	Ilumina Paired-e 301	854542						
Cj-RM35925_S23_L001_R1_001.fastq.gz	Ilumina Paired-e 301	859505						
Gj-RM35925_S23_L001_R1_001.fastq.gz	Ilumina Paired-e 301	749965						
Gj-RM35925_S23_L001_R2_001.fastq.gz	Ilumina Paired-e 301	870000						
Cj-RM35925_S23_L001_R2_001.fastq.gz	Ilumina Paired-e 301	770000						
Bowtie 2 - 16:00 - Dec 14, 2018								
Unaligned_Reads_71.fq.gz	Ilumina Unpaired 301	28175						
AnannotatedGenome Contig 71	1618194	1581295						
Welvet 1 - 15:38 - Dec 14, 2018								
Velvet Cj-RM35925_S23-1643_L001_Rx_001 - 15:32 - Dec 14, 2018								
Unaligned_Reads_1	454 301							
Contig 2	166101		166101 1 166					
Contig 3	36487		36487 1 36					
Contig 4	87616		87616 1 87					
Contig 5	241784		241784 1 241					
Contig 6	191924		191924 1 191					
Contig 7	38682		38682 1 38					
Contig 8	159227		159227 1 159					
Contig 9	47436	25686 1	47436 1 47					

There is a properties tab that lists all of the relevant details for the currently selected job, including input parameters, total assembled reads, average contig length, total contig length and even N50 to give you an idea of the quality of genome assemblies. A popup menu lets you quickly switch between different jobs.

#### Coverage Tab

Assembly projects now have a **Coverage** report that allows you to select a reference sequence and directly compare multiple reference assemblies. You can compare different

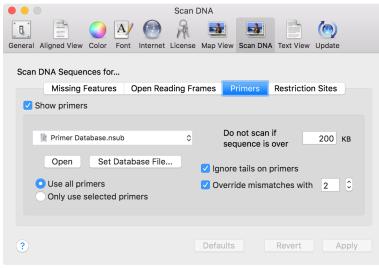
sequencing datasets assembled against the same reference sequence with expression level comparison. Now that you can store multiple assembly jobs in a single Assembly Project, then you can also directly compare multiple runs of the same dataset against a single reference sequence to compare which is the best assembly.

#### **Assembly Problems**

Align to Reference now has a **Problems** tab that helps you identify mis-assembled genomic sequences based on excessive mismatches, discontinuous reads and other common problems. The primary aim of this tab is to help you identify potential assembly issues with genomic sequencing projects. You can use Align to Reference to assemble several million reads against a typical small genome (practically, up to ~10 Mbp or so) and the tab will report those regions that are most likely to be misassembled.

#### **Scan for Primers**

MacVector now automatically scans DNA sequences for primer binding sites whenever you open a DNA sequence, in addition to the Restriction Enzymes, Open Reading Frames and Missing Features introduced in previous versions. By default, the collection of primers in the /MacVector/Subsequences/PrimerDatabase.nsub file are used, but you can use any .nsub file you like, configured using the updated MacVector | Preferences | Scan DNA preference settings.



You can right-click on any faded feature to add it to your sequence, or select all of them and add with a single mouse-click.

#### Use of APIKeys with Entrez and BLAST

The NCBI have introduced a new system for accessing the Entrez database over the internet. They now throttle general accesses to the Entrez server to no more than 3 calls per second per IP address. This might seem a lot, but if you want to download a list of 100 sequences, this can really slow things down, especially if you are sharing a public IP address behind a company or university firewall and others are using MacVector. They have introduced a new "APIKey" concept – you can request one of these from the NCBI for your own personal use and that increases the permitted rate to 10 calls per second and

is independent of the IP address. Thus, you can share a public IP address with colleagues and still each have the full 10 calls per second access.

Here's what the NCBI have to say about APIKeys;

https://ncbiinsights.ncbi.nlm.nih.gov/2017/11/02/new-api-keys-for-the-e-utilities/

Once you get one, you can tell MacVector to use it using the MacVector | Preferences | Internet tab;

			Inter	net				
		A/ 🥘	R (	-3	-b	NAMESCA	$(\bigcirc)$	
General Aligned	d View Color	Font Intern	et License	Map View	Scan DNA	Text View	Update	
Blast Server								
Current:	https://bla	st.ncbi.nlm.r	ih.gov/Bla	st.cgi				0
		Replace	Ad	d	Remove			
Server:	blast.ncbi.r	lm nih gov						
Server.	Diastinobili	inn.nin.gov						
Port:	443	Path:	/Blast.cgi					
Entrez Serve	er							
	_							_
E-Utilities API Key:								
				Defaults		Revert	Ap	ply

#### How Do I?

There's a brand new **How Do I** menu that shows common workflows with simple step by step guides and short videos.

Every tool dialog now has direct access to a simple video/tutorial on how to use it.

#### **Miscellaneous Enhancements**

You can now copy and paste features between sequence documents. Simply select the feature in the Features tab of one sequence, **Edit** | **Copy**, then switch to another sequence Features tab and choose **Edit** | **Paste**.

You can now select a short sequence in a Read in the Align To Reference **Editor** tab, right-click to bring up a context sensitive menu, and then select all of the other aligned reads that contain that a sequence. It's a great way of selecting all the reads with a specific SNP(s). Combine it with selecting the pairs for each read with NGS data and saving the selected reads and you have a very powerful way of analyzing SNPs or resolving repeat sequences in genomic assemblies.

Some of the nomenclature and interactions in the Primer3 dialog box have been cleaned up to simplify its use.

You can now add sequences to the *Align To Reference* window from multi-sequence GenBank files.

GenBank multi-sequence files now open by default as individual sequences.

When results are generated for the Bowtie **Coverage** tab, the name of each feature is now displayed exactly as per the label used in the **Map** tab allowing you more control over exactly how the text appears.

There is now an option in the **Database** | **Auto-annotate Sequence...** interface to automatically "fix" CDS features. This will change the stop location if the match creates a longer or shorter open reading frame and will also change the contents of any

/translation= qualifier to reflect the new translated coding sequence.

The *Align to Reference* algorithm is now multi-threaded. You will see a significant speed up, particularly if you are using larger *Sensitivity* values, though we still recommend keeping *Sensitivity* to a low (<4) value and the *Hash Value* to the maximum possible value for optimal performance when searching through large amounts of NGS data.

#### **Bug Fixes**

The plain text output from the **Analyze** | **Restriction Enzymes** function has been reworked to correct some problems introduced in MacVector 15 where the reported restriction enzyme cut sites were not correctly aligned.

Some inconsistencies in outlining similarities in the Multiple Sequence Alignment **Picture** tab have been resolved.

Some problems in the *Align to Reference* consensus calculation when aligning reads that wrap around circular sequences has been resolved.

Some strand-specific quirks with the Auto Annotation function have been resolved.

Network Licensing: Some issues with MacVector not cleanly releasing Sassafras KeyServer licenses have been cleaned up. In addition, the error messages reported with KeyServer are now more informative.

The scale bar has been restored to the *Phylogram* view in the *Phylogeny* window.

## Support information

For assistance with MacVector, please contact your local MacVector, Inc office. You will need a current MacVector maintenance contract to be eligible for technical support other than for basic installation or licensing problems. New sales of MacVector include 12 months of support that also entitles you to any upgrades to MacVector released during the maintenance period.

#### USA

Toll Free: (866) 338 0222

Telephone: (919) 303 7450

Fax: (919) 303 7449

E-mail: support@macvector.com

#### Europe

Telephone: + 44 (0)1223 410552

Fax: +44 (0)1223 709429

E-mail: support@macvector.com

#### Worldwide

Telephone: + 1 919 303 7450

E-mail: support@macvector.com

When contacting Customer Support with a technical problem, please be prepared to give your product serial number as well as a detailed description of your problem and any error messages you encounter. Visit the MacVector Web site for details of any available updates, and any relevant information that could not be added to these release notes in time for publication:

http://www.macvector.com

Maclector Inc.

Software for Scientists