## Sequence Analysis In VMD 1.8.7



#### Kirby Vandivort Senior Research Programmer

# Sequence and Structure Data Support

- PDB, ASTRAL, entire Swiss-Prot Database
- 100,000+ environmental sequences of Ribosomal RNA
- Sequence Data from BLAST, ALN, FASTA, Nexus, PIR, PHY, ALI
- Growth Temperature Metadata from the Prokaryotic Growth Temperature Database
- Taxonomic Sequence Information from the Comparative RNA Website

# Sequence and Structure Data Support (continued)

#### Secondary Structure Prediction via PSIPRED

🔻 Edit Sequence	e Information $ = \Box \times $
Sequence Name:	Myb.lepra3
Source Organism:	Mycobacterium leprae
Common Name:	
Sequence Type	RNA 🚄
EC Number:	
EC Description:	
Description:	Mycobacterium leprae str. LTB
Data Sources:	
Lineage:	Bacteria
	Actinobacteria
	Actinobacteridae Actinomycetales
	Corynebacterineae
	Myćobacteriaceae
Notes	ed the secondary structure below.
	Z
Secondary Structu	
Predict	
	OK Cancel

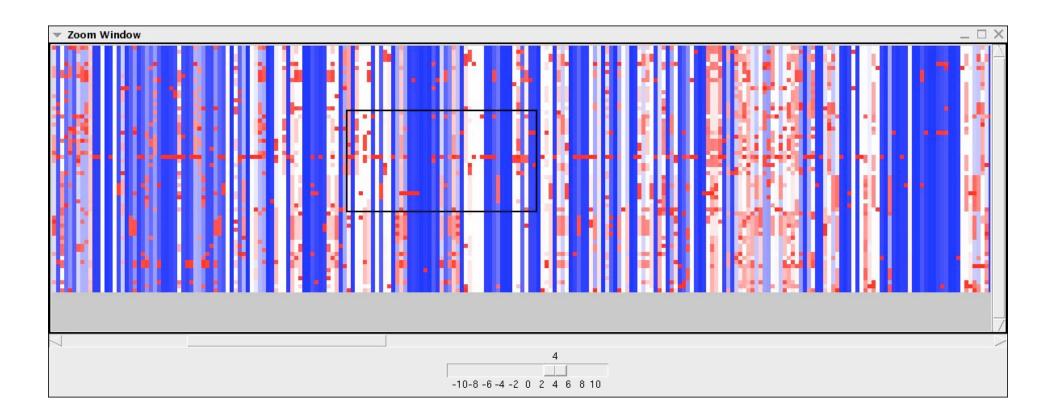
### Sequence Window

#### So Many Sequences.. So Little Screen Space..

▼ untitled.multiseq _ □ ×											
<u>F</u> ile Edit Sear	ch Tools	Options View				Help					
Sequence Name	2	20	230	240	250	260					
C.botulinD	ri 235 U (	C G G C C A C	AUUGGAAC	UGAGACACGG	UCCAGACUUAAO	CAGUGGGGA					
🗌 Pir.marina	ri 242 C	ССGUСUС	ACUGGGAC	U G A G A 🛄 A C 🛄 G	С С С А G А С 🗛 U А 🛄 Ф	I C A G U C G 🗛 G A					
Rps.palus7	ri 215 U (	САССАС	AUUGGGAC	UGAGACACGG	CCCAAACUUAAG	CAGUGGGA					
Zym.mobilis.3	ri 243 U (	САССАС	ACUGGGAC	UGAGACACGG	C C C A G A C U U A A G	CAGUGGGA					
] T.cuprina	ri 239 C (	САССАС	ACUGGGUC	UGAGACACGG	C C C A G A C U U A A G	CAGUGGGA					
] Nis.gonorr	ri 243 U (	СССС 🛯 АС	ACUGGGAC	UGAGACACGG	С С N А G А С U U А А С	I C A G U G G G N N					
] Nan.exeden	<b>ri</b> 243 U (	CAGUCAC	ACUGGAAC	UGAGACACGG	U C C A G A C U U A A G	CAGUGGGA					
] Cam.coli83	ri 240 U (	CAGUCAC	ACUGGAAC	UGAGACACGG	U C C A G A C U U A A G	I C A G U 🗛 G G A					
] Hlb.nemstr	ri 242 A	CGGACAC	ACUGGAAC	UGAGACACGG	U C C A G A C U U A A G	CAGUAGGA					
] WIn.succi2	ri 241 U (	C G G A C A C	ACUGGAAC	UGAGACACGG	U C C A G A C U U A A G	CAGUAGGA					
] Arm.hydrop	ri 242 U (	САССАС	ACUGGAAC	UGAGACACGG	U C C A G A C U U A A G	CAGUGGGA					
] Bor.garin5	ri 243 N (	GGGUGGN	NCUNGACO	U G A G A <mark>N</mark> A <mark>U</mark> G A	NACAGNCACCGO	CAGCNAAGA					
] Trp.pallidum.rrnA	ri 242 C	CGGACAC	ACUGGGAC	UGAGAUACGG	CCCAGACUUAAG	CAGCUAAGA					
] Fer.island	ri 244 C (	CGGCCAC	AAGGGCAC	UGAGACACGG	G C C C U A C U U A A C	I C A G U G G G G G					
] Tt.maritima.2.GEN	ri 242 U (	CGGCCAC	AGGGGCAC	UGAGACACGG		CAGUGGGGA					
] D.radiodu2	ri 217 C			UGAGACACGG	G U C C C A C U U A A C	C A G U U A G G A					
] Ap.pernix1	ri 237 G (	CCCCCAG	AUGGGCAC	UGAGACAAGG	G C C C A G G C U A A 🕻	CAGGCGCGA					
] Dco.mobili	ri 237 G	CCCCCAG	AUGGGCAC	UGAGACA <mark>A</mark> GG	G C C A G G C U A A C	CAGGCGCGA					
] Adi.brierl	ri 216 G	CCCCCAG	UUGGGCAC	UGAGACA <mark>AU</mark> G	G C C A G G C U A A C	CAGGCGCGA					
] Adi.infer2	ri 221 G	CCCCCAG	UUGGGCAC	UGAGACA <mark>A</mark> GG	G C C C A G G C U A A C	CAGGCGCGA					
] Sul.acald4	ri 237 G	С С 🚺 С С А G	UUGGGNNN		G C C C A G G C U A A C	CAGGCGCGA					
] Sul.solfa6	ri 226 G (	CCCCCAG	UUGGGCAC	UGAGACAAGG	G C C C A G G C U A A C	CAGGCGCGA					
] Tmf.penden	ri 237 G	C C C C <mark>G A</mark> G	AUGGGCAC	UGAGACA <mark>A</mark> GG	G C C C A G G C U A A C	CAGGGGCGA					
] Thp.tenax	ri 237 G	ССССБАБ	AUGGGCAC	UGAGACA <mark>A</mark> GG	G C C C A G G C U A A G	CAGGCGCGA					
C.symbiosum.A	ri 239 G	C C C <mark>G G A</mark> G	AUGGGUAC			CAGGCGAGA					
Arq.fulqidus	ri 237 G	C C C <mark>G G A</mark> G	AUGGACCC	UGAGACACGG		CAGGCGCGA					
						>					
Percent GC Content: 55.3389											

### Zoom Window

#### Quickly find regions of interest



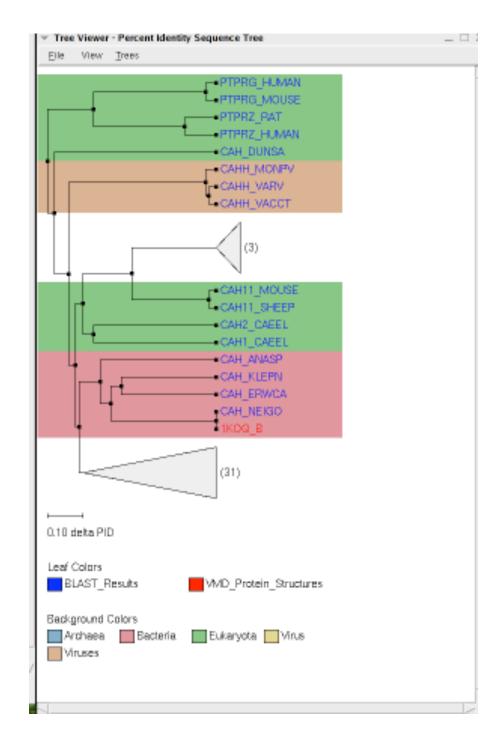
## Nucleic Acids

- Automatic differentiation between DNA/RNA/Protein sequences
- QR now works with nucleic acid structures and sequences (as well as protein structures and sequences)

	🔻 unti	tled.mu	ltiseq																	24	_		×
	<u>F</u> ile	Edit	Search	Tools		0	pti	ons	3	۷	'iev	V										Hel	p
	Sequen	ce Name	9		1									10									$ \Delta $
			Structures								333 												1
		_BR	vri	2	Α	Q	1	Т	G	R	Ρ	Е	W	1	W	L	А	L	G	Т	А	LI	
			Structures	1																			
		_ADNA	vri		A																	A	
		_BDNA	vri																			G	
	□ 1ffz_	_A	vri	2063	U	U	С	С	A	G	U	G	С	G	G	A	G	U	С	U	G	G	
n																							
																							-
					_		_	_	_	-	_	_	_	_	_	_	_	_	_	_	_	125	11
																						and the second	-

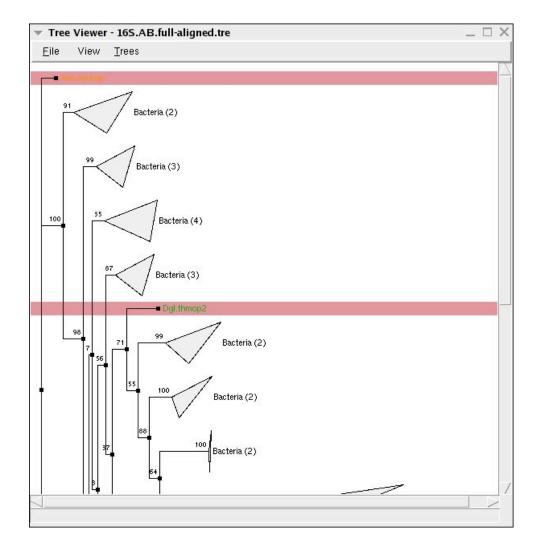
## Phylogenetic Trees

 MultiSeq tool to show
evolutionary
relationship
among
sequences



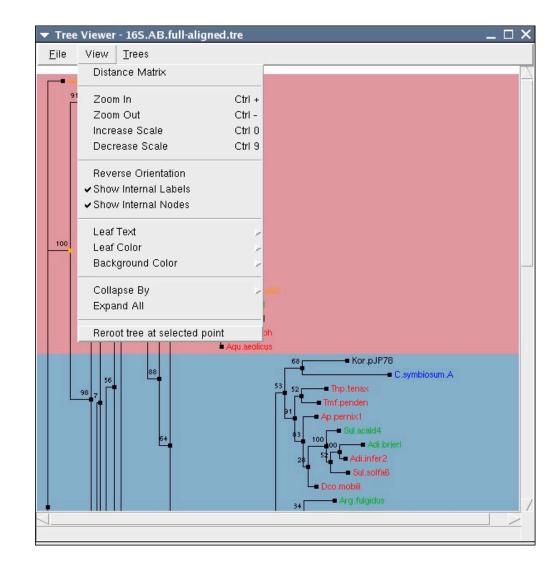
## Phylogenetic Trees

 Automatic collapsibility by taxonomy level, Temp. class, MultiSeq group



## Phylogenetic Trees

- Ability to "reroot" tree at a point of your choosing
- Distance-based tree from only well-aligned columns



## Advanced Feature - Scripting

- Use MultiSeq features in your own scripts and plugins
  - Shannon entropy/mutual information for sequence alignment
  - Combining alignments based on reference alignment
  - Structural Contact Information (Q)
  - Phylogenetic Trees manipulations, balance