

BIONUMERICS Tutorial:

Creating a minimum spanning tree based on MLVA repeat numbers

1 Aim

In this tutorial we will create a minimum spanning tree based on MLVA repeat numbers. We will also see how we can alter the layout of the minimum spanning tree and how to export the picture to use it in a publication, presentation, etc.

2 Preparing the database

- 1. Create a new database (see tutorial "Creating a new database") or open an existing database.
- 2. Import the MLVA repeat numbers from the example text file MLVA repeat numbers.txt as described in the tutorial: "Importing MLVA repeat numbers from a text file". This sample file contains repeat numbers for about 500 strains.

After import the Main window should look like Figure 1.

MLVA text import - BioNumerics		= U X
File Edit Database Analysis Scripts Window Help		
Experiment types	Database entries	Comparisons
👌 + 💾 ⊗ 🕄 Ē□ 🔽 ↑ ↓ «AllExe	第 + 🖄 😣 🗄 🧠 《All Entries》 🛛 🙂	+ 💾 🛞 🕄 🔂 🔍 (All Comparisons>
# Name Type 🗸	Key Modified date City 🗨 1	Name Modified date Level 🔻
Character types	ISO001 2020-04-09 09:39:30 Barcelona •	^
	ISO002 2020-04-09 09:39:30 Barcelona •	
	V ISO003 2020-04-09 09:39:30 Geneva •	
	ISO004 2020-04-09 09:39:30 Lyon •	< > >
	✓ ISO005 2020-04-09 09:39:30 Lyon •	
Entry fields Database design	ISO006 2020-04-09 09:39:30 Lyon •	Identification projects Decision networks
	✓ ISO007 2020-04-09 09:39:30 Lyon •	
	VISO008 2020-04-09 09:39:30 Lyon •	
Name Field type	ISO009 2020-04-09 09:39:30 Lyon •	Name Modified date 🗸
At City Fixed	VISO010 2020-04-09 09:39:30 Lyon •	
	VISO011 2020-04-09 09:39:30 Lyon •	
	VISO012 2020-04-09 09:39:30 Bordeaux •	
· · · · · · · · · · · · · · · · · · ·	Valencia 2020-04-09 09:39:30 Valencia •	
Fingerprint files Power assemblies Appolations	Valencia 2020-04-09 09:39:30 Valencia	Alignments BLAST projects Chromosome comparisons
	VISO015 2020-04-09 09:39:30 Lyon •	
I + I ⊗ E, E ⊂. <all files="" fingerprint=""></all>	ISO016 2020-04-09 09:39:30 Barcelona •	+ 🗁 🛞 🖏 🛛 🖬 🔽 - «All Alignments»
File name Experiment type Link	V ISO017 2020-04-09 09:39:30 Paris •	Name Modified date 🗨
	V ISO018 2020-04-09 09:39:30 Madrid •	
	V ISO019 2020-04-09 09:39:30 Madrid •	
	V ISO020 2020-04-09 09:39:30 Madrid •	
(V ISO021 2020-04-09 09:39:30 Madrid •	
	100000 2000 00.00.00 United	
Detailance MUXA load interest (Defaultilance) Entrine Landade 500 Minute 500 Salad		d ===k===

Figure 1: The Main window after import of the data.

The character data is stored in the character type MLVA.

Double-click on the experiment MLVA in the Experiment types panel, select Settings > General settings... (11), select the Experiment card tab and make sure the representation is to List. Close the two windows.

4. Click on a green colored dot in the *Experiment presence* panel of the *Main* window to open the experiment card for an entry.

The imported repeat numbers are displayed in the experiment card next to the corresponding locus name (see Figure 2).

/ ISO001			
Character	Value	Mapping	-
VNTR1	1	<+>	
VNTR2	47	<+>	
VNTR3	25	<+>	
VNTR4	7	<+>	

Figure 2: The experiment card.

- 5. Close the experiment card by clicking in the left upper corner of the card.
- 6. Right-click on the **City** information field in the *Database entries* panel of the *Main* window and choose *Field properties* from the floating menu (see Figure 3).

Data	base entries			
₽ ₊ +] + 🖻		<a>All Entries	ن ا
	Кеу		City	v 1
	ISO001	2020-04-09 09:39:30	Barcelona	Field properties
	ISO002	2020-04-09 09:39:30	Barcelona	Disable bigblighted field
	ISO003	2020-04-09 09:39:30	Geneva _	
	ISO004	2020-04-09 09:39:30	Lyon	Sort by field
	ISO005 2020-04-09 09:39:30		Lyon	Sort by field (reverse)
	ISO006	2020-04-09 09:39:30	Lyon	Soft by field (reverse)
	ISO007	2020-04-09 09:39:30	Lyon	Sort by field (Numerical)
	ISO008	2020-04-09 09:39:30	Lyon	Sort by field (Numerical, reverse)
	ISO009	2020-04-09 09:39:30	Lyon –	· · · · ·
	ISO010	2020-04-09 09:39:30	Lyon	Add information field
	ISO011	2020-04-09 09:39:30	Lyon	Delete information field
	IS0012	2020-04-09 09:39:30	Bordeaux	
	IS0013	2020-04-09 09:39:30	Valencia	Use as display field
	ISO014	2020-04-09 09:39:30	Valencia	Edit field in selection Ctrl+M
	ISO015	2020-04-09 09:39:30	Lyon	Copy field in selection
	IS0016	2020-04-09 09:39:30	Barcelona	
	ISO017	2020-04-09 09:39:30	Paris	

Figure 3: Field properties.

- 7. Press < Add all> to create all existing states for the City field. Confirm the action.
- 8. Check *Use colors* to display a specific color code for each field state (see Figure 4).
- 9. Press < OK > to accept the new settings.

The Database entries panel is updated (see Figure 5).



It is also possible to create groups based on the field content in the *Comparison* window with *Groups* > *Create groups from database field*.

Properties for field 'CITY'				?	×
Name: City		Changes: A	Allow		\sim
Field storage type: Optimized for spee	d	Restrict content	to states		
Content type: lext	~	Opacity:	-		
Field states Barcelona Geneva Lyon Bordeaux Valencia Paris Madrid	Add Add all Remove Remove all Edit Sort	Default content:	(regular expr. Level assi Summary Calcul); setting: ation Ok Cance	

Figure 4: The Database field properties dialog box.

MLVA text import - BioNumerics				– 🗆 ×
File Edit Database Analysis Scripts Window Help				
Experiment types	Database entries			Comparisons
		All Entries>	(5	+ 🖓 🛞 🖳 🗗 🔽 (All Comparisons)
		City.		
* Name Type	Key Modified date	City	· · ·	Name Modified date Level
L 1 MLVA Character types	L ISO001 2020-04-09 09:39	:30 Barcelona	•	
	L ISO002 2020-04-09 09:39	:30 Barcelona	•	
	L ISO003 2020-04-09 09:39	.30 Geneva		
	2020-04-09 09:39	Lyon Lyon		< >
Entry fields Database design	L ISO005 2020-04-09 09:39	Lyon		Mentification projects
	2020-04-09 09:39	Lyon		
+ 🖆 ⊗ 🗟 Ēli 🔽 ↑ ↓ <ali entry="" fields=""></ali>	2020-04-09 09:39	Lyon		2 + 1 × 1 × 1 × 1 × 1 × 1 × 1 × 1 × 1 × 1
Name Field type	2020-04-09 09:39	Lyon		Name Modified date 🗸 🗸
Als Chy Eived	2020-04-09 09:39	Lyon		
	ISO010 2020-04-09 09:39	20 Lyon		
	2020-04-09 09:39	20 Bordeaux		
×	ISO012 2020-04-09 09:39	30 Valencia		
	ISO014 2020-04-09 09:39	30 Valencia		
Fingerprint files Power assemblies Annotations	ISO015 2020-04-09 09:39	30 I von		Alignments BLAST projects Chromosome comparisons
□ + P ⊗ 民 日 、 <all files="" fingerprint=""></all>	ISO016 2020-04-09 09:39	30 Barcelona		+ 🖻 🛇 🗛 🗗 🔽 (All Alignments)
Sile name	ISO017 2020-04-09 09:39	:30 Paris		Name Medified date
rite name experiment type enix	ISO018 2020-04-09 09:39	30 Madrid		Name Modified date
	ISO019 2020-04-09 09:39	:30 Madrid		
	ISO020 2020-04-09 09:39	:30 Madrid		
	ISO021 2020-04-09 09:39	30 Madrid		
< > >	D 100000 00 00 00 00 00 00	20 Hadrid	· · · · · · · · · · · · · · · · · · ·	
Database: MLVA text import (_DefaultUser_) Entries: Loaded=500, View=500, Selected	I=0 1 experiments C:\Users\Public\Docu	ments\BioNumerics\Data BN8\MLVA text in	mport This is a time limited package valid until :	1020-12-30

Figure 5: The *Main* window.

3 Comparison window

- 1. Click somewhere in the *Database entries* panel of the *Main* window to make it the active panel, and select all entries using *Edit* > *Select all* (Ctrl+A).
- Highlight the *Comparisons* panel in the *Main* window and select *Edit* > *Create new object...* (+) to create a new comparison for the selected entries.
- 3. Click on the onext to the experiment name **MLVA** in the *Experiments* panel and select *Characters* > *Show values* () to display the repeat numbers in the *Experiment data* panel (see Figure 6).

Comparison												-	. 🗆	×
File Edit Layout Groups Clustering Stat	tistics Fingerprints Characters Sequence	TrendDat	a F	eadSe	s Spec	tra Con	nposit	e Window Help						
86 ≒×1°;	A Ca MLVA	Ē		b , 1		<u>ħ</u> ,								
Experiments														
<all experiment="" types=""></all>	🔍 Dendrogram	Experi					Infor	nation fields				Similarities		
Name A			123	1233	AIC (13 ec	R,	↓ \$↑ ↓₩₩	↑ ↓			i 123		
🔘 📑 MLVA 🚽	^	🛄 ML	VA											
<	~	묥		8	1R4									
		¥	₹	¥	₹	- 1		Key	City	Modified date	-			
Analyses Local composite datasets		1	47	25	7		~	150001	Barcelona	2020-04-09 09:39:30				^
P ^A ⊗		1	47	25	8		~	IS0002	Barcelona	2020-04-09 09:39:30				
Name		2	13	34	15		~	ISO003	Geneva	2020-04-09 09:39:30				
		2	5	4	8		~	ISO004	Lyon	2020-04-09 09:39:30				
		2	5	15	8		~	ISO005	Lyon	2020-04-09 09:39:30				
		2	10	15	9		~	ISO006	Lyon	2020-04-09 09:39:30				
		2	8	19	8		~	ISO007	Lyon	2020-04-09 09:39:30				
(C		4	8	36	10		~	ISO008	Lyon	2020-04-09 09:39:30				
Groups		4	8	38	8		~	ISO009	Lyon	2020-04-09 09:39:30				
		2	6	21	11		~	ISO010	Lyon	2020-04-09 09:39:30				
Size Name		2	7	26	9		~	IS0011	Lyon	2020-04-09 09:39:30				
oizo manto		3	8	27	11		~	IS0012	Bordeaux	2020-04-09 09:39:30				
		1	10	17	10		~	ISO013	Valencia	2020-04-09 09:39:30				
		7	2	30	9		~	IS0014	Valencia	2020-04-09 09:39:30				
					0			1			>	1		2.14

Figure 6: The *Comparison* window.

4 Advanced clustering window

A minimum spanning tree in BIONUMERICS is calculated in the *Advanced cluster analysis* window. This window can be launched from the *Comparison* window:

1. In the *Comparison* window select *Clustering* > *Calculate* > *Advanced cluster analysis...* or press the 🖆 button and select *Advanced cluster analysis* to launch the *Create network* wizard.

Due to the arbirariness of the repeat numbers, the similarity coefficient for clustering MLVA repeat numbers is the categorical coefficient. The categorical coefficient compares the repeat numbers to see if they are the same or different but does not quantify the difference. The predefined template *MST for categorical data* uses the categorical coefficient for the calculation of the similarity matrix, and will calculate a standard minimum spanning tree with single and double locus variance priority rules.

Specify an analysis name (for example MLVA1), make sure MLVA is selected, select MST for categorical data, and press < Next>.



To view and modify the settings of a selected template check the option *Modify template settings for new analysis*.

The *Advanced cluster analysis* window pops up. The *Network panel* displays the minimum spanning tree, the upper right panel (*Entry list*) displays the entries that are present in the tree. The *Cluster analysis method panel* displays the settings used, in this example the priority rules that result in the displayed network.

The colors of the comparison groups are automatically shown as node colors, but since no groups were defined in our comparison, no nodes are colored in our example. The coloring can very easily be changed to a field state grouping defined in the *Main* window:

3. Press 🖽 or choose *Display* > *Display settings* to open the *Display settings* dialog box.

4. In the *Node colors tab* select the **City** from the list and press <*OK*>.

The node colors are updated according to the city names. A node or branch can be selected by clicking on them. To select several nodes/branches hold the **Shift**-key, or click and hold down the left mouse button and drag the mouse pointer over the nodes to be selected.

5. Hold the Shift-key and select a few nodes.

The *Selection entry list panel* displays the entries currently selected on the network. The *Entry data panel* displays the character data for the selected entries.

- 6. The zoom slider on the left always further zooming in or out on the network. The zoom slider on top adjusts the size of the nodes.
- 7. Select *Display* > *Zoom to fit* or press S to optimize the view of the tree.
- 8. Press 🖽 or choose *Display* > *Display settings* to open the *Display settings* dialog box again.
- 9. Uncheck the option *Separate entries* in the *Node colors tab*.
- 10. Press < OK > to apply the new settings.

The Advanced cluster analysis window should now look like Figure 7.



Figure 7: The Advanced cluster analysis window.

- 11. The image can be exported with *File* > *Export image*.
- 12. Close the Advanced cluster analysis window and Comparison window with File > Exit.