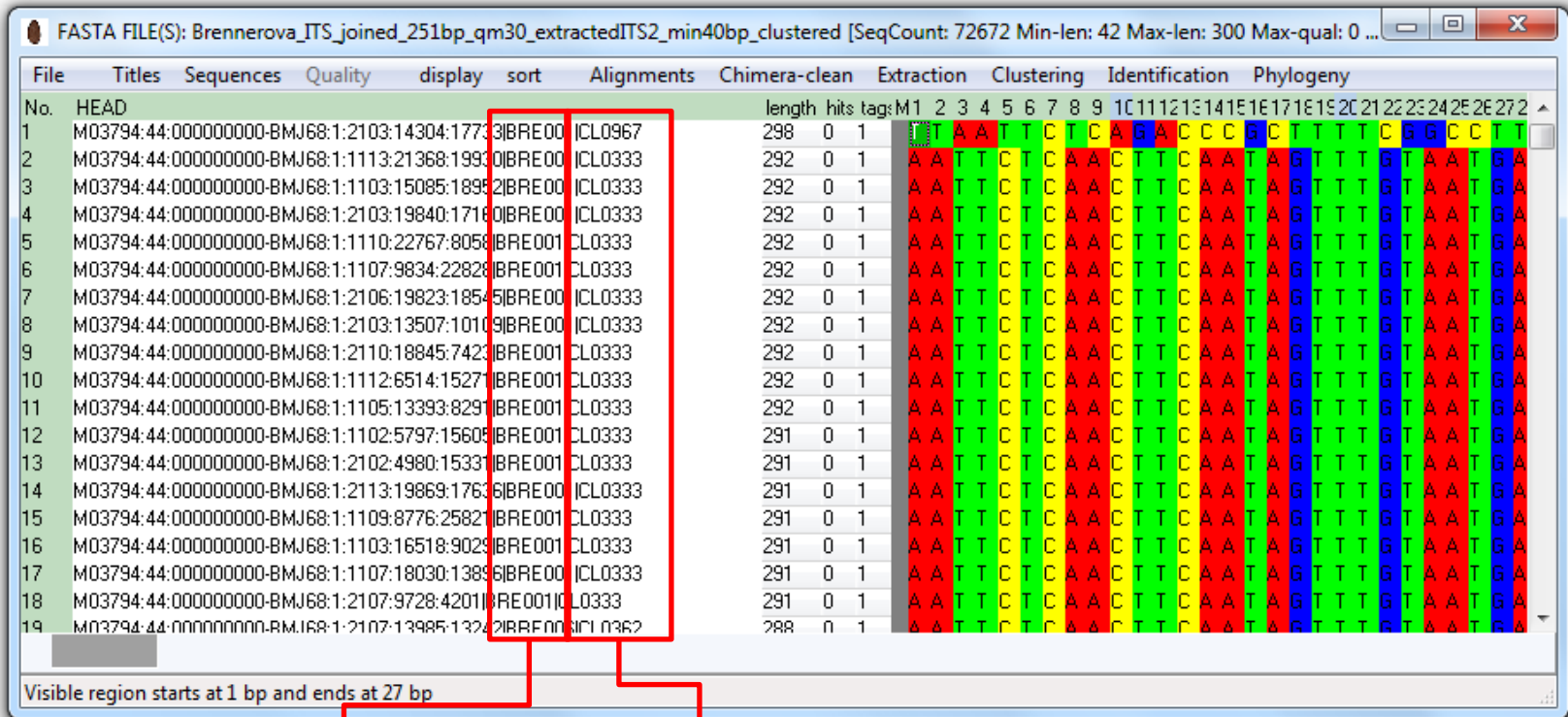


# How to estimate alpha diversity indices in SEED2

Step 1 – open FASTA file with sample and cluster names in titles (e.g.: Page 29 in "SEED2\_TUTORIAL\_16S.pdf" or Page 31 in "SEED2\_TUTORIAL\_ITS.pdf")



Sample name

Cluster name

## Step 2 – group by samples

The screenshot shows a bioinformatics software window titled "FASTA FILE(S): Brennerova\_ITS\_joined\_251bp\_qm30\_extractedITS2\_min40bp\_clustered [SeqCount: 72672 Min-len: 42 Max-len: 300 Max-qual: 0 ...". The main window displays a list of sequences with their headers and a corresponding alignment view. A context menu is open over the list, with "Group by titles" highlighted. A red box highlights this menu item, and a red arrow points to the "Group by titles" dialog box.

The "Group by titles" dialog box has two radio buttons: "Grouping by position in title" (unselected) and "Grouping by text inside two strings in title" (selected). The "Grouping by text inside two strings in title" option has input fields for "Starting pattern:" and "Ending pattern:", and a checkbox for "Include searched text" which is unchecked. A red box highlights these input fields. Below the dialog, a "Group" button is highlighted with a red box, and a red arrow points from it to the "0 of selected sequences" window.

The "0 of selected sequences" window shows a table with the following data:

No.	Group name:	
1	BRE004	25097
2	BRE001	16250
3	BRE002	14290
4	BRE006	7983
5	BRE005	6261
6	BRE003	2791
	no hit	0

## Step 2 – subsample sample to same amount of sequences

The process is shown in four stages:

- Initial state:** 0 sequences selected. A context menu is open over the table, with "Select all cells" highlighted.
- Selection:** 72672 of selected sequences. The "Randomly select sequences" option is highlighted in the context menu.
- Configuration:** A dialog box "Set random sequence selection:" is shown with "Set number of sequences:" set to 2791. The "OK" button is highlighted.
- Final state:** 72672 of selected sequences. The table shows the resulting counts for each group, and an "Info" dialog box displays "Random selection is finished." with an "OK" button.

No.	Group name:	Count
1	BRE004	25097
2	BRE001	16250
3	BRE002	14290
4	BRE006	7983
5	BRE005	6261
6	BRE003	2791
	no hit	0

### Step 3 – create OUT table with subsampled samples

16746 of selected sequences

Selected cells (rows) Divided dataset Analysis Table

BACK

Edit sequences  
Quality edit

Search and group by nucleotide motives 2D

**Group by titles 2D**

No.		
1		2791 (2503)
2		2791 (1625)
3		2791 (1429)
4	BRE006	2791 (7983)
5	BRE005	2791 (6261)
6	BRE003	2791 (2791)
	no hit	0

Group by titles

Grouping by position in title Position: 1 Length: 7

Grouping by text inside two strings in title

Starting pattern: ICL

Ending pattern:

include searched text

**Group**

Index Initiation ...  
Indexing...  
indexing done  
clusters 6

100%

### Step 4 – estimate (alpha) diversity indices

0 of selected sequences

Selected cells (rows) Divided dataset Analysis Table

BACK Show Selected

**Diversity indices**

Rarefaction calculator

No.	Group name:	BRE004	BRE001	BRE002	BRE006	BRE005	BRE003	no hit
1	ICL0001	166	174	137	155	62	119	0
2	ICL0002	164	17	38	114	199	30	0
3	ICL0003	120	11	20	7	379	9	0
4	ICL0005	49	97	97	83	62	86	0
5	ICL0007	51	130	75	86	55	75	0
6	ICL0004	57	98	49	111	56	96	0
7	ICL0006	71	122	86	63	55	56	0
8	ICL0009	41	98	68	58	71	69	0
9	ICL0008	65	76	80	71	41	59	0
10	ICL0010	35	59	64	39	48	73	0
11	ICL0012	36	38	60	64	24	59	0
12	ICL0011	28	64	69	21	6	70	0
13	ICL0014	45	40	30	38	42	61	0
14	ICL0016	35	22	49	25	11	89	0
15	ICL0013	60	40	43	49	17	15	0
16	ICL0015	32	51	28	37	50	7	0
17	ICL0017	14	34	51	31	23	42	0

table (Diversity indices)

Copy to clipboard Save as...

100%

	BRE004	BRE001	BRE002	BRE006	BRE005	BRE003	no hit
Shannon-Wiener Diversity Index	4.98689210553124	4.80650832655423	5.09164822529368	4.868594			
Shannon Entropy	7.19456451009823	6.93432572671129	7.34569564458231	7.02389679385546	6.817412		
Species Richness (S)	389	341	384	303	372	374	0
Total Abundance	2791	2791	2791	2791	2791	2791	0
Simpson Diversity Index	0.0156115507169035	0.0174899331564412	0.01255442937907				
Evenness	0.836224659418282	0.824177843994054	0.855646794492722				0.852086
Species Richness - 80% diversity	119	94	131	96	103	122	0
Chao-1	508.095238095238	419.241935483871	420.575	417.51724137931	509.459016393443	423.5696	