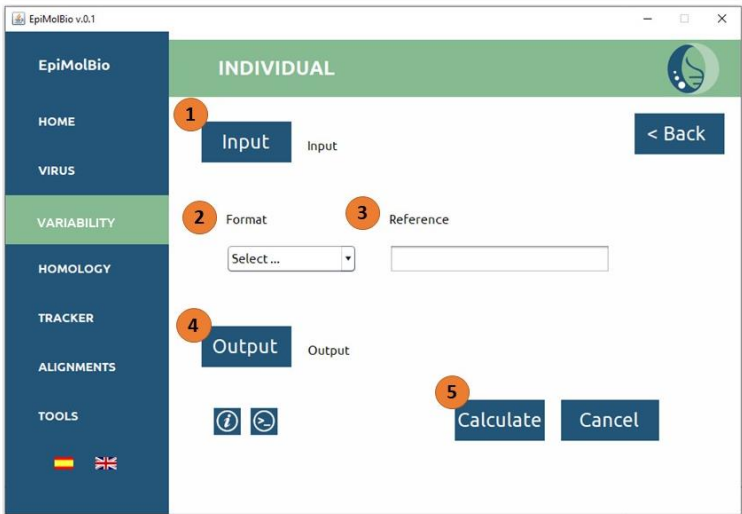


Variability Individual Conservation

It allows determining the level of conservation of sequences of interest by reporting the most prevalent residue and its percentage, and generating consensus sequences.



- 1 Select the input folder. It should only contain .fasta files of aligned NT or AA sequences to be analyzed (either one file or multiple files).
- 2 Choose the output format by selecting either **List** (employs automatic filtering >75%) or **Table** (employs automatic filtering of 100%).
- 3 Enter the reference sequence in letters without line breaks
- 4 Select the output folder and enter the file name in the following format: /NAME.html.
- 5 Click Calculate.

List:

Variability Conservation Individual List			
PR_01_AE.fasta			
CONSENSUS	PGITLWQRPLVTYKIGGQLKEALLDTGADDTVLEDINLPGKWKPKMIGGIGGFIVRQYDQILIEICGKIAGTTLVGPTPVNIIGRNMLTGCTLNF		
	Position	Residues	Total Positions
	P1	P(99.896%)	26638
	Q2	Q(99.782%)	26649
	V3	I(99.856%)	26831
	T4	T(99.858%)	26816
	L5	L(99.888%)	26780
	W6	W(99.929%)	26836
	Q7	Q(99.751%)	26536

Displays the generated consensus, the positions containing a conserved residue with a frequency greater than 75%, the residue with its colored frequency according to the color code, and the number of valid sequences for that position.

Table:

Variability Conservation Individual Table																	
File	Reference	P	Q	V	T	L	W	Q	R	P	L	V	T	I	K	I	G
	Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
PR_01_AE.fasta	Residue	P	Q	I	T	L	W	Q	R	P	L	V	T	V	K	I	G
	Conservation	99.896	99.782	99.858	99.858	99.888	99.929	99.751	99.922	99.955	87.446	99.512	95.028	56.099	95.188	88.392	72.550
	Number of Sequences	26838	26649	26831	26816	26780	26836	26536	26792	26613	25952	26416	26469	26150	26270	26206	25636
PR_02_AG.fasta	Residue	P	Q	I	T	L	W	Q	R	P	L	V	T	V	R	I	G
	Conservation	99.948	99.819	99.529	99.728	99.958	99.875	99.838	99.895	99.947	83.181	96.532	89.274	91.362	59.803	85.337	70.800
	Number of Sequences	9571	9416	9557	9561	9560	9575	9248	9557	9351	9186	9313	9295	9308	8916	9289	9233

This output displays the reference sequence and the position of each residue, followed by the analysis of each file with the most frequent residue por each podition, colored according to the color code, its conservation frequency, and the number of valid sequences for each position.

More information about this function can be found in the User Manual, section:

II. VARIABILITY II.2 CONSERVATION II.2.A) INDIVIDUAL