

Multiple Sequence Alignment Project

Damien GARCIA – Florian ECHELARD

January 2023

Contents

1	Traces generation	2
2	Alignment	2
3	MSA quality assessment	2

List of Figures

List of Tables

1	blableblibloblubly	2
---	------------------------------	---

Table 1: blablebliblioblubly

Required for simple generation		×	×	×	×	×							
Required for complex generation	×	×	×	×	×	×		×					×
Expression	<	(min	-	max)	+	event	X	val		...	>

1 Traces generation

The first section of the project aims to generate sequences called traces which are composed of tops and events. Events can be anchors, meaning they are shared by all generated traces and only exist once for each trace, or simple events that depending on the generation parameters, either exist in a trace or not, and could in some cases have repetitions. For easier further analysis of the traces, tops are represented by a dot (“.”), events always start by a full case “E” for anchors and lower case “e” for other events.

In order to generate traces, the program relies on parameters files providing a RegEx-like¹ sequence, the number of traces to generate and the maximal size of the traces to generate.

Listing 1: Parameters files example

```

1 | # GENERATION PARAMETERS
2 | expression=(2-3)E1(10-12)E9(1-4)E3
3 | number_of_traces=20
4 | maximum_length=100

```

2 Alignment

3 MSA quality assessment

¹Regular Expression