

Fasta format required by Projector:

Projector uses the absolute coordinates in its output files to indicate the position of the predicted genes and features. In order to do this, the input files have to come in a variant of the default fasta format which requires a header line of the format below.

Header line:

- >name start_position-end_position orientation

where

- name is the name of the sequence (example: Mm.X13235.5)
- start_position is an integer which is the position of the first character in the sequence (example: 100) and its value has to be smaller than that of the end_position
- end_position is an integer which is the position of the last character in the sequence (example: 737 i.e. the sequence is 737-100+1 = 638 nucleotides long)
- orientation can be either 'forward' or 'reverse'**depending on strand which is to be analysed for genes**. Note that the value of the orientation in the header line **does not** indicate the orientation of the sequence (gggaatg....)&as the fasta file should **always give the sequence of the forward strand**.
- the fields in the header line have to be tab-delimited

The DNA sequence:

- has to consist of A, C, G and T only
- always has to be the sequence of the forward strand

Example:

```
>Mm.X13235.5 100-737 forward
gggaatgaagtttttctgcaggatttaaattgtggtctttaagagacaccgcatgcaaaga
atagctggggcttgctagccaatgaaaacattcagattccaatgacgcatccttttttct
ccaccccttccaagaccggttcggaaccccgccctaacgctctagttttcaaccagg
tccgcagaaggcctatttaaggacgattgctgtctccctgctgtcataacatgtctg
gacgtggcaagggtggtaaaggccttgggaaaggcggcgctaagcgccaccgtaagggtc
tccgcgataacatccaggcatcaccaagcctgccatccgccgctggcccgcgcgggg
gagtgaagcgcctcctccggcctcatctacgaggagaccgcggtgtgctgaagggttcc
tggaacagtgatccgagcgcctcacctacacggagcacgccaagcgcaagaccgtca
ccgcatggacgtggtctacgctcaagcgccaggccgacctctctacggattcggcg
gttaatcgactaacaacgattttccactgtcaacaaaaggcccttttcagggccacca
caaattcctagaaggagttgttcacttacccaagcct
```