

De novo assembly exercise

Your job in this exercise is to de novo assemble the ‘genome’ from which the ‘reads’ below derive.

Rules/info:

- (*) Like real sequencing data, these reads contain errors. The error rate is ~2%
- (*) These are single-end 11-base reads
- (*) The average coverage is ~6x
- (*) You’re not allowed to google the answer
- (*) You can use your computers (i.e. word processors or text editors) or paper and whatever strategy you want to do the assembly...

_gew_kjinns	get_djinns_	l_get_djinn
n_outdjinn_	_can_outdji	en_if_they_
djinns_that	_j_will_get	en_if_they_
_are_djinns	tdjinn_them	y_are_djinn
will_get_dd	djinns_i_wi	an_outdjinn
en_if_they_	re_djinns_i	ns_i_willpg
_if_they_ar	e_djinns_i_	t_can_outdj
_will_get_d	ven_if_they	an_outdjinn
hat_can_out	t_can_outdj	inns_i_will
will_get_dj	_ehat_can_o	jinnns_i_wil
_djinns_tha	_djinns_tha	that_gan_ou
_if_they_ar	t_djinns_th	utdjinn_the

This exercise was inspired by and created using C. Titus Brown’s Assembly Exercise:
<http://ivory.idyll.org/blog/the-assembly-exercise.html>