

paired-end reads pipeline (Illumina)

■ Feature requests

U

umberar7

Oct 2010

Hi,

I was wondering how far fetched it would be to add a feature(s) to analyze paired end reads from Illumina platform. There are more and more publication coming out which are sequencing 16S using illumina platform. It would be interesting if Mothur could have a pipeline to analyze this type of data. I realize that we could still use Mothur to analyze single end reads but currently not much has been developed for paired-end reads. Any thoughts?

Thanks

♡

💬

10.4k views 



pschloss

Oct 2010

This is definitely in our scopes. So far I'm only aware of Rob Knight's PNAS study where they weren't able to pair the reads because the quality was too low. Are you aware of other studies? I think a big problem is going to be if the reads don't form contigs then what to do. What type of features are you looking for that would be unique to Illumina? My feeling is that once you have a contig you're good to go into mothur, but I could be missing something.

Pat

♡

💬

U

umberar7

Oct 2010

Hi Pat,

there is another paper where they developed their own pipeline to analyze paired end . It was published in PLoS ONE in august 2010 vol.5 by R. Hummelen. One of the authors was Gregory Gloor (I believe he developed the code) and he is in the process of publishing a methodology paper on analyzing paired ends.

You can find the paper at <http://arxiv.org/abs/1007.5075>. And yes the biggest problem is matching an overlap region between the two reads, other than that I think mothur could handle the rest. Let me know what you think. We are very interested in comparing 454 results to an Illumina run but need a way to analyze the paired-end reads.

Thanks

♡

💬

1 month later

G

ggloor

Nov 2010

We have now published the methods paper and have an automated pipeline to take paired-end Illumina reads all the way from the reads to the overlap to the ISU and OTU sequences. Feel free to contact me or my team members for more information

gg

♡

💬

2 years later

V

vcorby

Jan 2013

I see this is a rather old post, but I'm hoping for some insight...

The newest mothur announcement says

"...with the next release we should have a vetted SOP for analyzing Illumina data. Preliminary indications are that we get comparable sequence length to 454 that is just as good (or better) but a lot more reads."

I also get the feeling from this post that most people are using Illumina to do amplicon sequencing.

So, here's the big question - any idea of how to get diversity measures, etc. with shotgun Illumina data (i.e., not just 16S)? Some stuff on MG-RAST looks promising, but I like the straightforward nature of mothur.

Thanks for any insight.

Oct 2010

1 / 6

Oct 2010

Jan 2013

pschloss

Jan 2013

If you can make an rabund/sabund/list/shared file you can do any of the alpha or beta diversity metrics in mothur.

10.4k views

U

G

V

Reply

🔗 Related topics

Topic	Replies	Views	Activity
Single end read processing	3	653	Dec 2020
Processing paired end reads separately ■ Commands in mothur	3	1.7k	May 2016
Analysis of Illumina data - problem with make.contigs ■ Commands in mothur	10	9.9k	Sep 2014
Problems with dist.seqs and illumina reads ■ mothur bugs	1	2.5k	Jan 2014
AVITI vs. MiSeq	1	107	Aug 2024