From: **Lau, Matthew K.** matthewklau@fas.harvard.edu
Subject: Re: assemblies
Date: December 7, 2016 at 4:53 PM
To: Terrance Shea tshea@broadinstitute.org
Cc: Caroline Cusick ccusick@broadinstitute.org, Sarah Towey stowey@broadinstitute.org, jboch@broadinstitute.org

Hi Terry, thanks, I figured that this was likely the case. Thanks for doing this legwork.

To explain my interest a bit more, my inquiry is in part to "backup" this project, but it's also a part of another project on collecting data "provenance" for reproducible research. Here are a few papers (from other groups, not ours) that you might check out if you were interested:

http://bib.oxfordjournals.org/content/early/2016/03/23/bib.bbw020.full
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4248810/
http://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

Best,

Matt

> On Dec 7, 2016, at 4:13 PM, Terrance Shea <tshea@broadinstitute.org> wrote:
>
> Hi Matthew-
> Thanks for the heads-up on the assembly files being picked up.
>
> Unfortunately it is not a straightforward process with a single script that runs from end to end.  The pipeline calls a lot of different tools from various areas.  What I have done is attached the main log files which will point to some of the tools used or at least give a sense of how many different steps are involved in process from start to finish. The "a42" log described data preparation and assembly.  The "assembly.log" file is more specific to just the running of ALLPATHS-LG.  The "ap" log describes post-processing (including Pilon, though I had to run Pilon outside the pipeline to get it to complete due to memory limits).  And the "aa" log describes the GAEMR analysis done.

If you were to have to recreate the steps the easiest thing might be to just get in touch with us. It would be nearly impossible (or at least take a very long time) to try to piece everything together from these log files.

Terry

On Wed, Dec 7, 2016 at 3:31 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

> Hi Terry, a couple of (hopefully) quick things:
>
> 1) I don't think I ever emailed to let you know that I was done transferring the assemblies.
> 2) Do you have a script for all or any part of the assembly process? It would be good to have a record of that, also, it would be useful to replicate the process if needs/interests should arise.
>
> Thanks,
>
> Matt
>
>> On Nov 22, 2016, at 1:00 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:
>>
>> Great! Thanks Terry.
>>
>> Initiating transfer now…
>>
>> Best,
>>
>> Matt
>>
>>> On Nov 22, 2016, at 12:46 PM, Terrance Shea <tshea@broadinstitute.org> wrote:
>>>
>>> Hi Matt-
>>>
>>> These are all set. Let me know if you encounter any problems

with the ftp process.

Terry

On Tue, Nov 22, 2016 at 11:28 AM, Terrance Shea <tshea@broadinstitute.org> wrote:
No problem.  I will let you know once the tar jobs are complete.

On Tue, Nov 22, 2016 at 11:24 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

<SM-AJDMWa42.o8683339><SM-AJDMW.assembly.log><SM-AJDMWap42.o8741350><SM-AJDMWaa42.o8817692>