

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEII17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswwkd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366
[978-756-6165](tel:9787566165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles

Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicc
hio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
in
e is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
> wrote:

Hi
Jim,
just
follo
wing
up on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
the

sequence
labeling
issue?
Also,
is there
a good
solution to
removing
the
primers
from the
assemblies?

Best,
Matt

*Postdoctoral
Fellow*
Harvard
Forest
Harvard
University
324 N
Main
St.
Petersham,
MA
01366
[978-756-6165](tel:978-756-6165)

"Knowledge is
knowing a
tomato
is a
fruit.
Wisdom is
not
putting
it in a
fruit
salad."
—
Miles
Kingston

--
Jim
Bochicchio
Sr,
Product
Manager
Broad
Technology Labs
Broad
Institute
75 Ames
Street
Cambridge MA

02142,
USA

Phone: [617-714-8513](tel:617-714-8513)
Website : <https://www.broadinstitute.org/btl>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA
<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Jennifer M. Wineski
Senior Administrative Assistant
Broad Technology Labs

MIT-Broad Foundry

*Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411*

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

From: Lau, Matthew K. matthewklau@fas.harvard.edu
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: June 13, 2017 at 1:47 PM
To: Caroline Cusick ccusick@broadinstitute.org
Cc: Aaron M Ellison aellison@fas.harvard.edu, Jim Bochicchio jboch@broadinstitute.org



Hi Caroline, thanks for putting those on the FTP server. I'm grabbing those right now.

Best,

Matt

On Jun 13, 2017, at 9:43 AM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt and Aaron,

Terry has moved all the information to the FTP (we will still create the Aspera backup) and has provided the information on the new assemblies below:

The assembly data is available via <ftp://aadata@ftp.broadinstitute.org> with the following username and password, you should be able to access the directory /lau/ant:

- username **aadata**
- password **CILanalysis154**

Each directory will have 5 files. For example using this folder, SM-AJDMW/

SM-AJDMW/remove.txt
SM-AJDMW/assembly.agp
SM-AJDMW/scaffolds.fasta
SM-AJDMW/assembly.removed.fasta
SM-AJDMW/contigs.fasta

The contigs.fasta and scaffolds.fasta are the updated assembly files and the assembly.agp is the updated AGP corresponding to the updated fasta files. The "remove.txt" file is the table of directives instructing modify_assembly.py where to remove adapter sequence. Finally, assembly.removed.fasta is the fasta sequence of the removed adapter.

We're working on the methods and scripts, and I'll follow up with those as soon as we have them. Let me know if you have any questions or have any issues accessing the data.

thanks!
Caroline

On Wed, May 31, 2017 at 4:22 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Jim, thanks for the update.

Cheers,

Matt

On May 31, 2017, at 4:14 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Thank you!
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

I nank you very mucn,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Wednesday, May 24, 2017 10:39 AM

To: Jim Bochicchio <jboch@broadinstitute.org>

Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>

Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St.

324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this: "Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that such covered all the sites

make sure to confirm that we've removed all the sites
flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch
up with your data analysts yet? I have a meeting to
discuss the state of this project after lunch and it
would be good to have an idea of what you guys
are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a
fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just
finalizing a few things with our analyst
and I'll have an update for you
tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again
to make sure I don't fall of the
radar.

Thanks!

Matt

On May 10, 2017, at
10:13 AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please
keep me posted. A
quick question, if
there are primers still
in the assemblies,
what are the potential
issues with the
estimates of coverage,
GC content,
contaminant removal,

etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicchi
o
<jboch@broadinstitute.org
> wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this. I'm
going to
be out
of the
office
for the
next
few
days so
Carolin
e is
going to
keep an
eye on
this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM, Lau,
Matthew
K.
<matthew.klau@fas.harvard.edu>
wrote:

Hi Jim,
just
followi
ng up

on my
last
couple
of
emails
. Were
you
able to
find
out
anythi
ng
about
the
seque
nce
labelin
g
issue?
Also, is
there
a good
solutio
n to
removi
ng the
primer
s from
the
assem
blies?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvard
Forest
Harvard
Univers
ity
324 N
Main
St.
Petersh
am, MA
01366
978-
756-
6165

"Knowle
dge is
knowing
a
tomato
is a fruit.
Wisdom
is not
putting
it in a
fruit
salad."
— Miles
Kingston

--
Jim
Bochicchi
o
Sr,
Product
Manager
Broad
Technoln

Broad
Technology Labs
Broad
Institute
75 Ames
Street
Cambrid
ge MA
02142,
USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry

Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

From: Ellison, Aaron <aellison@fas.harvard.edu>
Subject: RE: Meeting request (was RE: ant sequences NCBI errors)
Date: June 13, 2017 at 9:48 AM
To: Lau, Matthew K. <matthewklau@fas.harvard.edu>

AE

Matt – as long as the ftp is clear, please download this all today.

Thanks,
Aaron

From: Caroline Cusick [mailto:ccusick@broadinstitute.org]
Sent: Tuesday, June 13, 2017 9:44 AM
To: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Jim Bochicchio <jboch@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron,

Terry has moved all the information to the FTP (we will still create the Aspera backup) and has provided the information on the new assemblies below:

The assembly data is available via <ftp://aadata@ftp.broadinstitute.org> with the following username and password, you should be able to access the directory /lau/ant:

- username aadata
- password CILanalysis154

Each directory will have 5 files. For example using this folder, SM-AJDMW/

SM-AJDMW/remove.txt
SM-AJDMW/assembly.agp
SM-AJDMW/scaffolds.fasta
SM-AJDMW/assembly.removed.fasta
SM-AJDMW/contigs.fasta

The contigs.fasta and scaffolds.fasta are the updated assembly files and the assembly.agp is the updated AGP corresponding to the updated fasta files. The "remove.txt" file is the table of directives instructing modify_assembly.py where to remove adapter sequence. Finally, assembly.removed.fasta is the fasta sequence of the removed adapter.

We're working on the methods and scripts, and I'll follow up with those as soon as we have them. Let me know if you have any questions or have any issues accessing the data.

thanks!
Caroline

On Wed, May 31, 2017 at 4:22 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Jim, thanks for the update.

Cheers,

Matt

On May 31, 2017, at 4:14 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Thank you!
Aaron

Aaron M. Ellison, Senior Research Fellow
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [mailto:jboch@broadinstitute.org]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Thursday, May 25, 2017 4:54 PM

To: Jennifer Wineski <jwineski@broadinstitute.org>

Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>

Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Wednesday, May 24, 2017 10:39 AM

To: Jim Bochicchio <jboch@broadinstitute.org>

Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswwkd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you

updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
224 Main St

324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicchio

<[jboch@broad
institute.
org](mailto:jboch@broadinstitute.org)>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
in
e is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<[matthe
wklau@
fas.harv
ard.edu](mailto:matthew.klau@fas.harvard.edu)
> wrote:

Hi
Jim,
just
follo
wing
up on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
the
seque
nce
labeli
ng
issue?
Also,
is
there

a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvar
d
Forest
Harvar
d
Univer
sity
324 N
Main
St.
Peters
ham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowl
edge is
knowin
g
a toma
to is a
fruit.
Wisdo
m
is not
putting
it in a
fruit
salad."
— Mile
s
Kingsto
n

--

Jim
Bochicc
hio
Sr,
Product
Manage
r
Broad
Technol
ogy Labs
Broad
Institute
75 Ames
Street
Cambrid
ge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Website

:

[https://
www.br
oadinsti
tute.org
/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski
Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry

Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

From: Ellison, Aaron aellison@fas.harvard.edu
Subject: RE: Meeting request (was RE: ant sequences NCBI errors)
Date: June 13, 2017 at 9:48 AM
To: Caroline Cusick ccusick@broadinstitute.org, Lau, Matthew K. matthewklau@fas.harvard.edu
Cc: Jim Bochicchio jboch@broadinstitute.org

AE

Thank you, Caroline!

Best,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Caroline Cusick [<mailto:ccusick@broadinstitute.org>]
Sent: Tuesday, June 13, 2017 9:44 AM
To: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Jim Bochicchio <jboch@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron,

Terry has moved all the information to the FTP (we will still create the Aspera backup) and has provided the information on the new assemblies below:

The assembly data is available via <ftp://aadata@ftp.broadinstitute.org> with the following username and password, you should be able to access the directory /lau/ant:

- username **aadata**
- password **CLanalysis154**

Each directory will have 5 files. For example using this folder, SM-AJDMW/

SM-AJDMW/remove.txt
SM-AJDMW/assembly.agp
SM-AJDMW/scaffolds.fasta
SM-AJDMW/assembly.removed.fasta
SM-AJDMW/contigs.fasta

The contigs.fasta and scaffolds.fasta are the updated assembly files and the assembly.agp is the updated AGP corresponding to the updated fasta files. The "remove.txt" file is the table of directives instructing modify_assembly.py where to remove adapter sequence. Finally, assembly.removed.fasta is the fasta sequence of the removed adapter.

We're working on the methods and scripts, and I'll follow up with those as soon as we have them. Let me know if you have any questions or have any issues accessing the data.

thanks!
Caroline

On Wed, May 31, 2017 at 4:22 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Jim, thanks for the update.

Cheers,

Matt

On May 31, 2017, at 4:14 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Thank you!
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street

324 NORTH MAIN STREET
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwscwckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pingin you again to make sure I don't fall of the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." - Billie

Matt Lau. — mmls
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bohicc
hio
<jbochc@broadcastinstitute.org>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
ine is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.k.lau@fas.harvard.edu>
> wrote:

Hi
Jim,
just
follo
wing
up on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
..

the
seque
nce
labeli
ng
issue?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvar
d
Forest
Harvar
d
Univer
sity
324 N
Main
St.
Peters
ham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowl
edge is
knowin
g
a toma
to is a
fruit.
Wisdo
m
is not
putting
it in a
fruit
salad."
— Mile
s
Kingsto
n

--
Jim
Bochicc
hio
Sr,
Product
Manage
r
Broad
Technol
ogy Labs
Broad
Institute
75 Ames
Street

Cambridge MA
02142,
USA

Phone:
[617-714-8513](tel:617-714-8513)

Website
:
[https://
www.br
oadinsti
tute.org
/btl](https://www.broadinstitute.org/btl)

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--

Lamifou M. Winkler

Jennifer M. Wineski
Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry

Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: [\(617\) 714-8411](tel:617-714-8411)

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

From: Caroline Cusick ccusick@broadinstitute.org
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: June 13, 2017 at 9:44 AM
To: Lau, Matthew K. matthewklau@fas.harvard.edu, Ellison, Aaron aellison@fas.harvard.edu
Cc: Jim Bochicchio jboch@broadinstitute.org

CC

Hi Matt and Aaron,

Terry has moved all the information to the FTP (we will still create the Aspera backup) and has provided the information on the new assemblies below:

The assembly data is available via <ftp://aadata@ftp.broadinstitute.org> with the following username and password, you should be able to access the directory /lau/ant:

- username **aadata**
- password **CILanalysis154**

Each directory will have 5 files. For example using this folder, SM-AJDMW/

SM-AJDMW/remove.txt
SM-AJDMW/assembly.agp
SM-AJDMW/scaffolds.fasta
SM-AJDMW/assembly.removed.fasta
SM-AJDMW/contigs.fasta

The contigs.fasta and scaffolds.fasta are the updated assembly files and the assembly.agp is the updated AGP corresponding to the updated fasta files. The "remove.txt" file is the table of directives instructing `modify_assembly.py` where to remove adapter sequence. Finally, assembly.removed.fasta is the fasta sequence of the removed adapter.

We're working on the methods and scripts, and I'll follow up with those as soon as we have them. Let me know if you have any questions or have any issues accessing the data.

thanks!
Caroline

On Wed, May 31, 2017 at 4:22 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:
Hi Jim, thanks for the update.

Cheers,

Matt

On May 31, 2017, at 4:14 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Thank you!
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Thursday, May 25, 2017 4:54 PM

To: Jennifer Wineski <jwineski@broadinstitute.org>

Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>

Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu

@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwscwckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall of the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicchi
o
<[jboch@
broadinst
itute.org](mailto:jboch@broadinstitute.org)
> wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this. I'm
going to
be out
of the
office
for the
next
few
days so
Carol
ine is
going to
keep an
eye on
this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM, Lau,
Matthew
K.
<[matthe
wklau@f
as.harvar
d.edu](mailto:matthew.klau@fas.harvard.edu)>
wrote:

Hi Jim,
just
followi
ng up
on my
last
couple
of
emails
More

. were
you
able to
find
out
anythi
ng
about
the
seque
nce
labelin
g
issue?
Also, is
there
a good
solutio
n to
removi
ng the
primer
s from
the
assem
blies?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvard
Forest
Harvard
Univers
ity
324 N
Main
St.
Petersh
am, MA
01366
978-
756-
6165

"Knowle
dge is
knowing
a tomat
o is a
fruit.
Wisdom
is not
putting
it in a
fruit
salad."
— Miles
Kingston

--
Jim
Bochicchi
o
Sr,
Product
Manager
Broad
Technolo
gy Labs
Broad
Institute
75 Ames

Street
Cambrid
ge MA
02142,
USA

Phone: 6
17-714-
8513
Website:
https://
www.bro
adinstitut
e.org/btl

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA
<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513
Website: https://www.broadinstitute.org/btl

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513
Website: https://www.broadinstitute.org/btl

--
Jennifer M. Wineski

Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry

Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA



From: Lau, Matthew K. matthewklau@fas.harvard.edu
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 31, 2017 at 4:22 PM
To: Jim Bochicchio jboch@broadinstitute.org
Cc: Jennifer Wineski jwineski@broadinstitute.org, Caroline Cusick ccusick@broadinstitute.org, Aaron M Ellison aellison@fas.harvard.edu

Hi Jim, thanks for the update.

Cheers,

Matt

On May 31, 2017, at 4:14 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Thank you!
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Wednesday, May 24, 2017 10:39 AM

To: Jim Bochicchio <jboch@broadinstitute.org>

Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>

Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this: "Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: (978) 756-6165

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall of the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 9, 2017, at 3:36 PM, Jim Bochicchio
<jbochicchio@broadinstitute.org>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
in
e is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
> wrote:

Hi
Jim,
just
follow
ing up
on my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
the
seque
nce
labeli
ng
issue?
Also,
is
there
a

good
solution to
removing
the
prime
rs
from
the
assemblies?

Best,

Matt

*Postdoctoral
Fellow*
Harvard
Forest
Harvard
University
324 N
Main
St.
Petersham,
MA
01366
978-
756-
6165

"Knowledge is
knowing a
tomato
is a
fruit.
Wisdom is
not
putting
it in a
fruit
salad."
—
Miles
Kingston

--
Jim
Bochicchio
Sr,
Product
Manager
Broad
Technology Labs
Broad
Institute
75 Ames
Street
Cambridge MA
02142,

USA

Phone:

617-

714-

8513

Website

:

<https://>

www.broadinstitute.org/btl

[oadinstitute.org](https://www.broadinstitute.org/btl)

[/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry

Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

From: Ellison, Aaron aellison@fas.harvard.edu
Subject: RE: Meeting request (was RE: ant sequences NCBI errors)
Date: May 31, 2017 at 4:14 PM
To: Jim Bochicchio jboch@broadinstitute.org
Cc: Lau, Matthew K. matthewklau@fas.harvard.edu, Jennifer Wineski jwineski@broadinstitute.org, Caroline Cusick ccusick@broadinstitute.org

Thank you!
Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Wednesday, May 31, 2017 3:56 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Wednesday, May 24, 2017 10:39 AM

To: Jim Bochicchio <jboch@broadinstitute.org>

Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>

Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were

adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:(978)756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thank!

THANKS:

Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.

<matthewklau@fas.harvard.edu>

wrote:

Hi Caroline, just pinging you again
to make sure I don't fall of the
radar.

Thanks!

Matt

On May 10, 2017, at
10:13 AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please
keep me posted. A
quick question, if
there are primers still
in the assemblies,
what are the
potential issues with
the estimates of
coverage, GC
content, contaminant
removal, etc.?

Thanks, Jim, have a
good few days out of
the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bohicc
hio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into

this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
ine is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
> wrote:

Hi
Jim,
just
follow
ing up
on my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
the
seque
nce
labeli
ng
issue?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime

primers
from
the
assemblies?

Best,

Matt

*Postdoctoral
Fellow*
Harvard
Forest
Harvard
University
324 N
Main
St.
Petersham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowledge is
knowing
a tomato is a
fruit.
Wisdom
is not
putting
it in a
fruit
salad."
— Miles
Kingston

--

Jim
Bochicchio
Sr,
Product
Manager

Broad
Technology Labs
Broad
Institute
75 Ames
Street
Cambridge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Website

:

[https://
www.br
oadinsti
tute.org
/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski
Senior Administrative Assistant

*Broad Technology Labs
MIT-Broad Foundry*

*Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: [\(617\) 714-8411](tel:6177148411)*

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

From: Jim Bochicchio jboch@broadinstitute.org
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 31, 2017 at 3:56 PM
To: Ellison, Aaron aellison@fas.harvard.edu
Cc: Lau, Matthew K. matthewklau@fas.harvard.edu, Jennifer Wineski jwineski@broadinstitute.org, Caroline Cusick ccusick@broadinstitute.org

JB

Hi Matt and Aaron

We wanted to send along a quick update. Our analyst Terry is going back through 5 of the assemblies you mentioned and working on trimming them. As for the other two assemblies, Terry is working with one of our engineers to put in place a pipeline they can run to look for the adapter sequencing and trim them since you didn't get an NCBI error report for those two that he can work off of.

Thanks
Jim

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks

Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.

Sent: Wednesday, May 24, 2017 10:39 AM

To: Jim Bochicchio <jboch@broadinstitute.org>

Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron
<aellison@fas.harvard.edu>

Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks

Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so

weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:(978)756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,

Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_
filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!

thanks,

Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

-- Matt Klingon

On May 11, 2017, at 4:55 PM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm
just finalizing a few things
with our analyst and I'll have
an update for you tomorrow.

thanks!

Caroline

On Thu, May 11, 2017 at
11:31 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, just ping
you again to make sure I
don't fall off the radar.

Thanks!

Matt

On May 10,
2017, at 10:13
AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline,
please keep me
posted. A quick
question, if there
are primers still
in the
assemblies, what
are the potential

issues with the
estimates of
coverage, GC
content,
contaminant
removal, etc.?

Thanks, Jim,
have a good few
days out of the
office.

Best,

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9,
2017,
at 3:36
PM,
Jim
Bochic
chio
<jboch@broa.dinstitute.org>
wrote:

Hi Matt

Our
analysts
are still

looking
into this.
I'm going
to be out
of the
office for
the next
few days
so
Caroline
is going
to keep
an eye
on this.

Thanks

Jim

On
Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klaus@fas.harvard.edu>
wrote:

Hi
Jim,
just
follow
ing
up on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anythi
ng
about
the
sequ
ence
labeli
ng

ing
issue
?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdoct
oral
Fellow*

Harvard
Forest

Harvard
Universit
y

324 N
Main St.

Petersha
m, MA
01366

[978-756-
6165](tel:978-756-6165)

"Knowled
ge is
knowing
a tomato
is a fruit.
Wisdom
is not
putting it
in a
fruit
salad." —
Miles
Kingston

--

Jim
Bochic
chio

Sr,
Product
Manag
er

Broad
Technol
ogy
Labs
Broad
Institute
75
Ames
Street
Cambri
dge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Websit
e:
[https://
www.br
oadinsti
tute.
org/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs

Broad Institute

75 Ames Street

Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs

Broad Institute

75 Ames Street

Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

Senior Administrative Assistant

Broad Technology Labs

MIT-Broad Foundry

*Broad Institute of MIT and Harvard
75 Ames Street*

*Office 8016A
Cambridge, MA 02142
Phone: [\(617\) 714-8411](tel:617-714-8411)*

Email: jwineski@broadinstitute.org

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

From: Jennifer Wineski jwineski@broadinstitute.org
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 5:03 PM
To: Ellison, Aaron aellison@fas.harvard.edu
Cc: Lau, Matthew K. matthewklau@fas.harvard.edu, Jim Bochicchio jboch@broadinstitute.org, Caroline Cusick ccusick@broadinstitute.org

JW

Thank you. I will send an invitation shortly!

On Thu, May 25, 2017 at 5:01 PM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Works for me, too.

We'll be there.

Thank you,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks

Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron
<aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks

Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,

Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_
filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged

that we've removed all the sites tagged
by NCBI.

let me know if you have any questions!

thanks,

Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau,
Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a
chance to catch up with your data
analysts yet? I have a meeting to
discuss the state of this project after
lunch and it would be good to have an
idea of what you guys are doing on
your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a
fruit, wisdom is not putting it in a fruit
salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!

Caroline

On Thu, May 11, 2017 at 11:31 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, just ping you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim,
have a good few
days out of the
office.

Best,

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9,
2017,
at 3:36
PM,
Jim
Bochic
chio
<jboch@broadsideminstitute.org>
wrote:

Hi Matt

Our
analysts
are still
looking
into this.
I'm going
to be out
of the
office for
the next
few days
so
Caroline

Caroline
is going
to keep
an eye
on this.

Thanks

Jim

On
Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
wrote:

Hi
Jim,
just
follow
ing
up on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anythi
ng
about
the
sequ
ence
labeli
ng
issue
?
Also,
is
there
a
good

soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdoct
oral
Fellow*

Harvard
Forest

Harvard
Universit
y

324 N
Main St.

Petersha
m, MA
01366

[978-756-
6165](tel:978-756-6165)

"Knowled
ge is
knowing
a tomato
is a fruit.
Wisdom
is not
putting it
in a
fruit
salad." —
Miles
Kingston

--

Jim
Bochicc
hio

Sr,
Product
Manag
er

Broad
Technol
ogy
Labs
Broad
Institute
75
Ames
Street
Cambri
dge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Websit
e:
[https://
www.br
oadinsti
tute.
org/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs

Broad Institute

75 Ames Street

Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs

Broad Institute

75 Ames Street

Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

Senior Administrative Assistant

Broad Technology Labs

MIT-Broad Foundry

*Broad Institute of MIT and Harvard
75 Ames Street*

Office 8016A

Cambridge, MA 02142

Phone: [\(617\) 714-8411](tel:6177148411)

Email: jwineski@broadinstitute.org

--

Jennifer M. Wineski

Senior Administrative Assistant

Broad Technology Labs

MIT-Broad Foundry

*Broad Institute of MIT and Harvard
75 Ames Street*

Office 8016A

Cambridge, MA 02142

Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

From: Ellison, Aaron aellison@fas.harvard.edu
Subject: RE: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 5:01 PM
To: Lau, Matthew K. matthewklau@fas.harvard.edu, Jennifer Wineski jwineski@broadinstitute.org
Cc: Jim Bochicchio jboch@broadinstitute.org, Caroline Cusick ccusick@broadinstitute.org

AE

Works for me, too.

We'll be there.

Thank you,
Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
[@AMaxEll17](https://twitter.com/AMaxEll17)
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Thursday, May 25, 2017 4:54 PM
To: Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Jim Bochicchio <jboch@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, *Senior Research Fellow*
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu
@AMaxEl17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcsckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM_AZXXM12). If the other six were returned with

(JIVEPZAAW): If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978.756.6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still

there are primers sum
in the assemblies,
what are the
potential issues with
the estimates of
coverage, GC content,
contaminant
removal, etc.?

Thanks, Jim, have a
good few days out of
the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochic
chio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analyst
s are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Carol
ine is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9

May 3,
2017 at
10:32
AM,
Lau,
Matthe
w K.

<matthew.k.lau@fas.harvard.edu>

> wrote:

Hi
Jim,
just
follow
ing up
on my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anythi
ng
about
the
seque
nce
labeli
ng
issue?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvar
d
Forest
Harvar
d
Univer
sity
324 N
Main

St.
Peters
ham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowl
edge is
knowin
g
a tomat
o is a
fruit.
Wisd
om
is not
putting
it in a
fruit
salad."
— Mile
s
Kingsto
n

--
Jim
Bochicc
hio
Sr,
Product
Manage
r
Broad
Technol
ogy Labs
Broad
Institute
75 Ames
Street
Cambrid
ge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)
Website
:
[https://
www.br
oadinstit
ute.org/
btl](https://www.broadinstitute.org/btl)

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA
<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski
*Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry*

*Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411*

Email: jwineski@broadinstitute.org

From: Lau, Matthew K. matthewklau@fas.harvard.edu
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 4:53 PM

To: Jennifer Wineski jwineski@broadinstitute.org

Cc: Jim Bochicchio jboch@broadinstitute.org, Aaron M Ellison aellison@fas.harvard.edu, Caroline Cusick ccusick@broadinstitute.org



Hi Jenn, thanks that works for me.

Best,

Matt

On May 25, 2017, at 4:52 PM, Jennifer Wineski <jwineski@broadinstitute.org> wrote:

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks

Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest

324 North Main Street
Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron
<aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks

Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswwkd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,

Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last

email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt".
Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!

thanks,

Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:(978)756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!

Caroline

On Thu, May 11, 2017 at 11:31 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017,
at 10:13 AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline,
please keep me
posted. A quick
question, if there
are primers still in
the assemblies,
what are the
potential issues
with the
estimates of
coverage, GC
content,
contaminant
removal, etc.?

Thanks, Jim,
have a good few
days out of the
office.

Best,

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a
tomato is a fruit. Wisdom is
not putting it in a

fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochic
chio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analysts
are still
looking
into this.
I'm going
to be out
of the
office for
the next
few days
so
Caroline
is going
to keep
an eye on
this.

Thanks

Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
> wrote:

Hi
Jim,
just
followi
ng up
on my
last
couple
of
emails
. Were
you
able
to find
out
anythi
ng
about
the
seque
nce
labelin
g
issue?
Also,
is
there
a
good
solutio
n to
remov
ing
the
primer
s from
the
assem
blies?

Best,

Matt

*Postdocto
ral Fellow*

Harvard
Forest

Harvard
University

324 N
Main St.

Petersha
m, MA
01366

[978-756-
6165](tel:978-756-6165)

"Knowled
ge is
knowing a
tomato is
a fruit.
Wisdom is
not putting
it in a
fruit
salad." —
Miles
Kingston

--

Jim
Bochicc
hio

Sr,
Product
Manage
r

Broad
Technol
ogy
Labs
Broad
Institute
75
Ames
Street
Cambri
dge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Website
:
<https://www.broadinstitute.org/btl>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

*Senior Administrative Assistant
Broad Technology Labs
MIT-Broad Foundry*

*Broad Institute of MIT and Harvard
75 Ames Street
Office 8016A
Cambridge, MA 02142
Phone: (617) 714-8411*

Email: jwineski@broadinstitute.org

From: Jennifer Wineski jwineski@broadinstitute.org
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 4:53 PM
To: Jim Bochicchio jboch@broadinstitute.org
Cc: Ellison, Aaron aellison@fas.harvard.edu, Lau, Matthew K. matthewklau@fas.harvard.edu, Caroline Cusick ccusick@broadinstitute.org

JW

Good afternoon Aaron and Matthew,

I can offer Monday, June 12th from 1pm -1:30pm for a meeting with Jim.

Will that work for your schedules as well?

Thank you,

Jenn

On Thu, May 25, 2017 at 4:41 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:
Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks

Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow

Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,

Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers.

Cusick,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt".
Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!

thanks,

Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!

Caroline

On Thu, May 11, 2017 at 11:31 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017,
at 10:13 AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline,
please keep me
posted. A quick
question, if there
are primers still in
the assemblies,
what are the
potential issues
with the
estimates of
coverage, GC
content,
contaminant
removal, etc.?

Thanks, Jim,
have a good few
days out of the
office.

Best,

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochic
chio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analysts
are still
looking
into this.
I'm going
to be out
of the
office for
the next
few days
so
Caroline
is going
to keep
an eye on
this.

Thanks

Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.
<matthew.klau@fas.harvard.edu>
> wrote:

Hi

Jim,
just
followi
ng up
on my
last
couple
of
emails
. Were
you
able
to find
out
anythi
ng
about
the
seque
nce
labelin
g
issue?
Also,
is
there
a
good
solutio
n to
remov
ing
the
primer
s from
the
assem
blies?

Best,

Matt

*Postdocto
ral Fellow*

Harvard
Forest

Harvard

University

324 N
Main St.

Petersha
m, MA
01366

[978-756-
6165](tel:978-756-6165)

"Knowled
ge is
knowing
a tomato
is a fruit.
Wisdom
is not
putting it
in a
fruit
salad." —
Miles
Kingston

--

Jim
Bochicc
hio

Sr,
Product
Manage
r

Broad
Technol
ogy
Labs
Broad
Institute
75
Ames
Street
Cambri
dge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Website

:

[https://
www.br
oadinstit
ute.or
g/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs

Broad Institute

75 Ames Street

Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jennifer M. Wineski

Senior Administrative Assistant

Broad Technology Labs

MIT-Broad Foundry

Broad Institute of MIT and Harvard

75 Ames Street

Office 8016A

Cambridge, MA 02142

Phone: (617) 714-8411

Email: jwineski@broadinstitute.org

From: Ellison, Aaron aellison@fas.harvard.edu
Subject: RE: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 4:43 PM
To: Jim Bochicchio jboch@broadinstitute.org, Jennifer Wineski jwineski@broadinstitute.org
Cc: Lau, Matthew K. matthewklau@fas.harvard.edu, Caroline Cusick ccusick@broadinstitute.org

AE

Thank you.

Aaron

Aaron M. Ellison, Senior Research Fellow
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Jim Bochicchio [<mailto:jboch@broadinstitute.org>]
Sent: Thursday, May 25, 2017 4:42 PM
To: Ellison, Aaron <aellison@fas.harvard.edu>; Jennifer Wineski <jwineski@broadinstitute.org>
Cc: Lau, Matthew K. <matthewklau@fas.harvard.edu>; Caroline Cusick <ccusick@broadinstitute.org>
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, Senior Research Fellow
Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA
aellison@fas.harvard.edu
@AMaxEll17
<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochnicchio <jbochn@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcdwckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the

make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at
10:13 AM, Lau,
Matthew K.
<matthewklau@fas.h

arvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicchio
<jboch@broadinstitute.org>
wrote:

Hi Matt

Our
analysts are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days

days
so
Carolin
e is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Lau,
Matthe
w K.

<matthew.klau@fas.harvard.edu>

> wrote:

Hi
Jim,
just
follow
ing up
on
my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anyth
ing
about
the
seque
nce
labeli
ng
issue?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime

rs
from
the
asse
mblie
s?

Best,

Matt

*Postdo
ctoral
Fellow*
Harvar
d
Forest
Harvar
d
Univer
sity
324 N
Main
St.
Peters
ham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowl
edge is
knowin
g
a toma
to is a
fruit.
Wisd
om
is not
putting
it in a
fruit
salad."
— Mile
s
Kingsto
n

--

Jim
Bochicc
hio
Sr,
Product
Manage
r
Broad
Technol
ogy Labs
Broad
Institute
75 Ames
Street
Cambrid

ge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Website
:
[https://
www.br
oadinsti
tute.org
/btl](https://www.broadinstitute.org/btl)

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)
Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

From: Jim Bochicchio jboch@broadinstitute.org
Subject: Re: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 4:42 PM
To: Ellison, Aaron aellison@fas.harvard.edu, Jennifer Wineski jwineski@broadinstitute.org
Cc: Lau, Matthew K. matthewklau@fas.harvard.edu, Caroline Cusick ccusick@broadinstitute.org

JB

Hi Aaron

We are adding Jenn to the email so she can assist in setting up this meeting.

Thanks
Jim

On Thu, May 25, 2017 at 9:20 AM, Ellison, Aaron <aellison@fas.harvard.edu> wrote:

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest

324 North Main Street

Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu

@AMaxEll17

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks

Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad (<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwcswckd>).

Per my request for the code, are your analysts starting to use these Workbench

methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 16, 2017, at 9:17 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,

Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_
filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by

specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!

thanks,

Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.

<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest

Harvard University

324 N Main St

Peterborough, MA 01366

Peter Sham, MA 01500

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM,
Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!

Caroline

On Thu, May 11, 2017 at 11:31 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10,
2017, at 10:13
AM, Lau,
Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline,
please keep me
posted. A quick
question, if there
are primers still
in the
assemblies, what
are the potential
issues with the
estimates of
coverage, GC
content,
contaminant
removal, etc.?

Thanks, Jim,
have a good few
days out of the
office.

Best,

Matt

Postdoctoral Fellow

Harvard Forest

Harvard University

324 N Main St.

Petersham, MA 01366

[978-756-6165](tel:978-756-6165)

"Knowledge is knowing

Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9,
2017,
at 3:36
PM,
Jim
Bochic
chio
<[jboch@broad
institute.org](mailto:jboch@broadinstitute.org)>
wrote:

Hi Matt

Our
analysts
are still
looking
into this.
I'm going
to be out
of the
office for
the next
few days
so
Caroline
is going
to keep
an eye
on this.

Thanks

Jim

On
Tue,
May 9,
2017 at
10:32
AM,
Lau,
...

Matthe
w K.

<matthewklau@fas.harvard.edu>

wrote:

Hi
Jim,
just
follow
ing up
on my
last
coupl
e of
email
s.
Were
you
able
to find
out
anythi
ng
about
the
seque
nce
labeli
ng
issue
?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdoctoral
Fellow*

Harvard
Forest

Harvard
University

324 N
Main St.

Petersham, MA
01366

[978-756-
6165](tel:978-756-6165)

"Knowledge is
knowing
a tomato
is a fruit.
Wisdom
is not
putting it
in a
fruit
salad." —
Miles
Kingston

--

Jim
Bochichio

Sr,
Product
Manager

Product

Broad
Technol
ogy
Labs
Broad
Institute
75
Ames
Street
Cambri
dge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)

Websit
e:
[https://
www.br
oadinsti
tute.
org/btl](https://www.broadinstitute.org/btl)

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--

Caroline Cusick

Product Coordinator

Broad Institute of MIT and Harvard

75 Ames St | Cambridge MA 02142 USA

--

Jim Bochicchio

Sr, Product Manager

Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: [617-714-8513](tel:617-714-8513)

Website: <https://www.broadinstitute.org/btl>

--

Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513

Website: <https://www.broadinstitute.org/btl>

From: Ellison, Aaron aellison@fas.harvard.edu
Subject: Meeting request (was RE: ant sequences NCBI errors)
Date: May 25, 2017 at 9:20 AM
To: Lau, Matthew K. matthewklau@fas.harvard.edu, Jim Bochicchio jboch@broadinstitute.org
Cc: Caroline Cusick ccusick@broadinstitute.org

AE

Dear Jim and Caroline,

Could you please check today with Jenn to set up this meeting for either June 8 afternoon or June 12 any time? I have a number of other meetings I need to schedule for those two days in Cambridge, and I want to ensure that that I (the PI on the ant project) can meet with you and Matt.

Thank you very much,

Aaron

Aaron M. Ellison, Senior Research Fellow

Harvard University, Harvard Forest
324 North Main Street
Petersham, Massachusetts 01366 USA

aellison@fas.harvard.edu
[@AMaxEll17](https://twitter.com/AMaxEll17)

<http://harvardforest.fas.harvard.edu/aaron-ellison/>

From: Lau, Matthew K.
Sent: Wednesday, May 24, 2017 10:39 AM
To: Jim Bochicchio <jboch@broadinstitute.org>
Cc: Caroline Cusick <ccusick@broadinstitute.org>; Ellison, Aaron <aellison@fas.harvard.edu>
Subject: Re: ant sequences NCBI errors

Hi Jim and Caroline, looking forward to getting something on the calendar.

Thanks!

Matt

On May 22, 2017, at 8:25 PM, Jim Bochicchio <jboch@broadinstitute.org> wrote:

Hi Matt

We'll check with Jenn our admin to set up a meeting. She's currently out of the office but should be back on Wednesday.

Thanks
Jim

On Thu, May 18, 2017 at 4:49 PM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, I know your analysts are still processing, but I figured it would be great to schedule an in-person meeting sometime within the next few weeks. These meetings can be hard to schedule for us in the summer time, so it's best to get something on the books now.

Does Thursday June 8 in the mid-to-late afternoon work for you? If not, does Monday June 12 (any time) work?

Also, I came across this presentation on Data Sharing at Broad
(<http://genomicinfo.broadinstitute.org/acton/attachment/13431/f-0132/1/-/-/-/-/ASHG16%20Workbench%20Presentation%20-%20Final.pdf?sid=TV2:QKwswckd>).

Per my request for the code, are your analysts starting to use these Workbench methods? It seems like it would make this a whole lot easier.

Thanks!

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing a tomato is a fruit. Wisdom is not putting it in a fruit salad." — Miles Kingston

On May 17, 2017, at 6:25 AM, Lau, Matthew K. <matthewklau@fas.harvard.edu> wrote:

Hi Caroline, great thanks. Also, please ask your analyst what the ETA is. I would also like to see the code for the pipeline. I had asked for this back in December, but Terry wasn't able to get it after the analysis had already been done.

Best,

Matt

Postdoctoral Research Fellow
Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:978-756-6165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kingston

On May 16, 2017, at 9:17 PM, Caroline Cusick <ccusick@broadinstitute.org> wrote:

Hi Matt,

Our analyst is going through the comments that you sent from NCBI now, and I will pass along any updates/questions he has during this and definitely keep you updated!

thanks,
Caroline

On Mon, May 15, 2017 at 8:44 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline and Jim, happy Monday!

What's the status on the assemblies? And per my last email, are you re-processing from the raw reads?

Cheers,

Matt

On May 12, 2017, at 12:14 PM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Great, thanks Caroline!

All 7 were flagged with errors, but only 5 have reports. I've concatenated the errors that I could get from NCBI (see attached). Each report starts with a header like this:
"Contamination.txt_AZXXM_filtered_scaffolds.txt". Let me know if you have any questions.

A question though. It seems like this means that there were adapters left in the sequences used for the assemblies. Are you going to go back to the raw sequences, remove them and then re-do the assemblies?

Just in case it's helpful to know now, your FTP server will work fine for the data transfer.

Cheers,

Matt

On May 12, 2017, at 11:58 AM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

We're going to go through and screen the assemblies for adapter, paying specific attention to the sites flagged by NCBI, and then set up a site for that data handoff. Once we get a sense of how long this will take, I can follow up with an ETA.

Our analyst had a question about whether all 7 assemblies were submitted to NCBI, or just this one (SM-AZXXM)? If the other six were returned with errors also, if you could pass that along and he will make sure to confirm that we've removed all the sites flagged by NCBI.

let me know if you have any questions!
thanks,
Caroline

On Fri, May 12, 2017 at 11:21 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Morning Caroline, have you had a chance to catch up with your data analysts yet? I have a meeting to discuss the state of this project after lunch and it would be good to have an idea of what you guys are doing on your end.

Thanks,

Matt

Postdoctoral Research Fellow

Harvard Forest
Harvard University
324 N Main St
Petersham, MA 01366

Office: [\(978\) 756-6165](tel:9787566165)

Knowledge is knowing that a tomato is a fruit, wisdom is not putting it in a fruit salad.

-- Miles Kington

On May 11, 2017, at 4:55 PM, Caroline Cusick
<ccusick@broadinstitute.org> wrote:

Hi Matt,

You are not forgotten! I'm just finalizing a few things with our analyst and I'll have an update for you tomorrow.

thanks!
Caroline

On Thu, May 11, 2017 at 11:31 AM,
Lau, Matthew K.
<matthewklau@fas.harvard.edu>
wrote:

Hi Caroline, just pinging you again to make sure I don't fall off the radar.

Thanks!

Matt

On May 10, 2017, at 10:13 AM, Lau, Matthew K.
<matthewklau@fas.harvard.edu> wrote:

Hi Caroline, please keep me posted. A quick question, if there are primers still in the assemblies, what are the potential issues with the estimates of coverage, GC content, contaminant removal, etc.?

Thanks, Jim, have a good few days out of the office.

Best,

Matt

Postdoctoral Fellow
Harvard Forest
Harvard University
324 N Main St.
Petersham, MA 01366
[978-756-6165](tel:978-756-6165)

"Knowledge is knowing
a tomato is a fruit. Wisdom
is not putting it in a
fruit salad." — Miles
Kingston

On May
9, 2017,
at 3:36
PM, Jim
Bochicchio
<jbochicchio@broadinstitute.org>
wrote:

Hi Matt

Our
analysts
are
still
looking
into
this.
I'm
going
to be
out of
the
office
for the
next
few
days
so
Caroline
is
going
to keep
an eye
on this.

Thanks
Jim

On Tue,
May 9,
2017 at
10:32
AM,
Jim

...
Matthe
w K.
<matthew.klaus@fas.harvard.edu>
> wrote:

Hi
Jim,
just
follow
ing up
on my
last
coupl
e of
email
s.
Were
you
able
to
find
out
anythi
ng
about
the
seque
nce
labeli
ng
issue?
Also,
is
there
a
good
soluti
on to
remo
ving
the
prime
rs
from
the
asse
mblie
s?

Best,

Matt

*Postdo
ctoral
Fellow
Harvar
d*

Forest
Harvard
University
324 N
Main
St.
Peters
ham,
MA
01366
[978-
756-
6165](tel:978-756-6165)

"Knowl
edge is
knowin
g
a tomat
o is a
fruit.
Wisd
om
is not
putting
it in a
fruit
salad."
— Mile
s
Kingsto
n

--
Jim
Bochic
chio
Sr,
Product
Manage
r
Broad
Technol
ogy Labs
Broad
Institute
75 Ames
Street
Cambrid
ge MA
02142,
USA

Phone:
[617-
714-
8513](tel:617-714-8513)
Website
:
[https://
www.br
oadinstit
ute.org/
h+l](https://www.broadinstitute.org/)

[BTI](#)

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

<ncbi_contam_screen.txt>

--
Caroline Cusick
Product Coordinator
Broad Institute of MIT and Harvard
75 Ames St | Cambridge MA 02142 USA

--
Jim Bochicchio
Sr, Product Manager
Broad Technology Labs
Broad Institute
75 Ames Street
Cambridge MA 02142, USA

Phone: 617-714-8513
Website: <https://www.broadinstitute.org/btl>

