Overview

Session One - Introducing Yabi (60 mins)
- Designing workflows
- Job Management
- Results and file management
- Command line access to Yabi

Session Two – Administering Yabi (60 mins)
- Workflow management
- Resources
- Users and Credentials
- Tools

Session Three – Installing Yabi (60 mins)
- Local installation for testing
- Deploying to the cloud
Session One - Introducing Yabi (60 mins)
- Designing workflows
- Job Management
- Results and file management
- Command line access to Yabi
Session One - Introducing Yabi

Workflows made easy.

Bioinformatics
Yabi does a brilliant job in bridging the gap between powerful tools and user friendliness. With its state-of-the-art web-based interface, it is a much needed - and very welcome - addition to the bioinformatics ecosystem.

— Tomas Di Domenico, Biocomputing Lab, Department of Biology, University of Padova

Parallel Analysis of Input files
Although the interface is linear in nature it is possible to achieve parallel branching over many files.
1.1 Design EBI workflow
1.2 Look at file system resources in Select File
1.2.1 Upload file
### 1.3 Tools

```plaintext
<table>
<thead>
<tr>
<th>Find tool:</th>
</tr>
</thead>
</table>

**Tools**

<table>
<thead>
<tr>
<th>ebi webservice</th>
<th>add</th>
</tr>
</thead>
<tbody>
<tr>
<td>clusterw2</td>
<td></td>
</tr>
<tr>
<td>fasta</td>
<td></td>
</tr>
<tr>
<td>fetchdb</td>
<td></td>
</tr>
<tr>
<td>ebi_ncbi_blast</td>
<td></td>
</tr>
<tr>
<td>ebi_wublast</td>
<td></td>
</tr>
</tbody>
</table>

**qlime**

<table>
<thead>
<tr>
<th>alpha_rarefaction</th>
<th>add</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta_diversity</td>
<td></td>
</tr>
<tr>
<td>beta_diversity_through_plots</td>
<td></td>
</tr>
<tr>
<td>check_id_map</td>
<td></td>
</tr>
<tr>
<td>jackknifed_beta_diversity</td>
<td></td>
</tr>
<tr>
<td>make_3d_plots</td>
<td></td>
</tr>
<tr>
<td>make_bootstrapped_tree</td>
<td></td>
</tr>
<tr>
<td>make_otu_heatmap_html</td>
<td></td>
</tr>
<tr>
<td>make_otu_network</td>
<td></td>
</tr>
<tr>
<td>per_library_stats</td>
<td></td>
</tr>
<tr>
<td>pick_otus_through_otu_table</td>
<td></td>
</tr>
<tr>
<td>principal_coordinates</td>
<td></td>
</tr>
<tr>
<td>split_libraries</td>
<td></td>
</tr>
<tr>
<td>summarize_taxa_through_plots</td>
<td></td>
</tr>
</tbody>
</table>

**select data**

| select_file | add |

**system**

| hostname (starcuster) | add |
```
1.3.1 Tool Description

The dbfetch services provide access to entries from various up-to-date biological databases using entry identifiers or accession numbers.

accepts: *.txt
outputs: *.fa
1.3.2 Inputs and outputs
1.3.2 Inputs and outputs

Options for 2 - ebi_ncbiblast

- program: blastx
  The BLAST program to be used for the Sequence Similarity Search.

- stype: DNA/RNA
  Indicates if the sequence is protein or DNA/RNA.

- sequence: 1 - select file
  Sequence
  - extension: bsl
  Converts the extension of the output from bsl to bsl

- database: UniProt Knowledgebase
  Database

- outfile: output
  Output file name to use, based on --sequence
1.3.3 Validation
1.4 Naming workflows

Options for 2 - ebi_ncbiblast

- program blastx
- stype DNA/RNA
  Indicates if the sequence is protein or DNA/RNA.
- sequence 1 - select file
  Converting the extension of the out option from b2s to b2s
- database UniProt Knowledgebase
- extension .txt
- outfile .output
  Output file name to use, based on --sequence
2.1 Running jobs
### 2.2 Job Status

<table>
<thead>
<tr>
<th>Job Status</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/20/12</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Account</th>
<th>sge</th>
</tr>
</thead>
<tbody>
<tr>
<td>CWD</td>
<td>/yabi/data/demo1/cjde6e050f-28af-45ea-be02-71791e6468d5/output</td>
</tr>
<tr>
<td>ENV_LIST</td>
<td>job_scripts/414</td>
</tr>
<tr>
<td>GID</td>
<td>1000</td>
</tr>
<tr>
<td>GROUP</td>
<td>Ubuntu</td>
</tr>
<tr>
<td>JOB_NAME</td>
<td>yabi-616</td>
</tr>
<tr>
<td>JOB_NUMBER</td>
<td>416</td>
</tr>
<tr>
<td>JOBSHARE</td>
<td>0</td>
</tr>
<tr>
<td>NOTIFY</td>
<td>Ubuntu@master</td>
</tr>
<tr>
<td>OWNER</td>
<td>Ubuntu</td>
</tr>
<tr>
<td>SCHEDULING_INFO</td>
<td>(Collecting of scheduler job information is turned off)</td>
</tr>
<tr>
<td>SCRIPT_FILE</td>
<td>/tmp/1001f0a2-3a6e-4897-a34e-0b90ed06651b.sh</td>
</tr>
<tr>
<td>SGE_O_HOME</td>
<td>/home/ubuntu</td>
</tr>
<tr>
<td>SGE_O_HOST</td>
<td>master</td>
</tr>
<tr>
<td>SGE_O_LOG_NAME</td>
<td>ubuntu</td>
</tr>
<tr>
<td>SGE_O_PATH</td>
<td>/as/local/yabi/tools/lab/tools/user/local/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/bin/input</td>
</tr>
<tr>
<td>Submission Time</td>
<td>Fri Oct 26 09:18:00 2012</td>
</tr>
<tr>
<td>Usage</td>
<td>cpu=0.000:00, mem=0.00000 GBs, ion=0.00000, vmem=N/A, maxvmem=N/A</td>
</tr>
</tbody>
</table>
2.3.1 Preview files
2.3.1 Preview files

```plaintext
-tz|BT7811|BT7811_T6OS3 SCF ubiquitin ligase, 60k component, putative OS=Ixodes scapularis GB=IscK_T6OS3 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDL CIRCAQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|C9MW92|C9MW92_9ACAS Putative uncharacterized protein OS=Ambyloma maculatum PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDL CIRCAQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
	+tz|C9MW92|C9MW92.ACT1 ACYP1002214 protein OS=Aegilops tauschii pisin GB=ACYP1002214 PE=4 SV=1 MEVESEDEVRPHFSTSECRKCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|C2AS11|C2AS11_TBICA Putative uncharacterized protein GLEAN_07016 OS=Triolium castaneum GB=GLEAN_07016 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|J2JY11|J2JY11.9CUCV Uncharacterized protein OS=Oedraectonus ponderosae PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|B0QX50|B0QX50.CULQ0 Ring-box protein 1A OS=Culex quinquefasciatus GB=CqipJ_CPI2J012886 PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|Q0K77|Q0K77_ABOGA AGAP0505292-PA OS=Anopheles gambiase GB=AGAP0505292 PE=4 SV=2 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|E3N9S4|E3N9S4_ABOGA Uncharacterized protein OS=Anopheles darlingi GB=AHM_26478 PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|H9K332|H9K332_API5 Uncharacterized protein OS=Apis mellifera GB=Rocia PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
	+tz|E3N9S4|E3N9S4_DAPP0 Putative uncharacterized protein OS=Sapia fulvescens GB=DAPPDRAFT_40730 PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
+tz|Q27519|Q27519_BOMTO Ring box protein OS=Bombus mori PE=4 SV=1 NAME=IscK_VDQ2QQELSTSSADGCEFRKVEKKAVALKWVNDVCAICRHIMDLIEIQQA
MQASATSECTVAVGCHAFHFSFCSIRKFLTRQCGFLHREMEFPKYGQH
```

11/20/12
2.4 Job re-use
3. Files Tab

<table>
<thead>
<tr>
<th>File Path</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qiime Tutorial (1)</td>
<td>37K</td>
</tr>
<tr>
<td>Qiime Tutorial (2)</td>
<td>37K</td>
</tr>
<tr>
<td>Qiime Tutorial (3)</td>
<td>37K</td>
</tr>
<tr>
<td>epimer</td>
<td>37K</td>
</tr>
<tr>
<td>epimer (1)</td>
<td>37K</td>
</tr>
<tr>
<td>AA123456.fa</td>
<td>0.61K</td>
</tr>
</tbody>
</table>
3.4 SFTP or S3 access
4. Qiime workflow example

Qiime Tutorial (3)

Tags:

- re-use

start

1. select file
2. check_id_map
3. split_libraries
4. pick_otus_through_otu_table
5. per_library_stats
6. make_otu_heatmap_html
7. make_otu_network
8. summarize_taxa_through_plots
9. alpha_rarefaction
10. beta_diversity_through_plots
11. jackknifed_beta_diversity
12. make_3d_plots

end

file outputs

53://ignored@yabi-sing.amazonaws.com/yabi-sing-prod/andrew/Qiime Tutorial (3)/12 - make_3d_plots

1. 3d_plots

2012-10-05 20:35:28

jar

unweighted_unifrac_pc_3D_PCoA_plots.html
4.1 Qiime on-line tutorial

Compute Alpha Diversity within the Samples and Generate Rarefaction Curves

Community ecologists typically describe the microbial diversity within their study. This diversity can be assessed within a sample (alpha diversity) or between a collection of samples (beta diversity). Here, we will determine the level of alpha diversity in our samples using a series of scripts from the QIIME pipeline. To perform this analysis, we will use the alpha_rarefaction.py workflow script. This script performs the following steps:

1. Generate rarefied OTU tables (for more information, refer to multiple_rarefactions.py)
2. Compute measures of alpha diversity for each rarefied OTU table (for more information, refer to alpha_diversity.py)
3. Collate alpha diversity results (for more information, refer to collate_alpha.py)
4. Generate alpha rarefaction plots (for more information, refer to make_rarefaction_plots.py)

Although we could run this workflow with the (sensible) default parameters, this provides an excellent opportunity to illustrate the use of custom parameters. To see what measures of alpha diversity will be computed by default, type:

\`alpha_diversity.py -h\`

You should see, among other information:

```
-m METRICS, --metrics=METRICS
   Alpha-diversity metric(s) to use. A comma-separated list should be provided when multiple metrics are specified. [default: PD_whole_tree,chaol,observed_species]
```

To also use the shannon index, create a custom parameters file by typing:

```
echo "alpha_diversity:metrics shannon,P0_whole_tree,chaol,observed_species" > alpha_params.txt
```

Then run the workflow, which requires the OTU table (-i) and phylogenetic tree (-t) from above, and the custom parameters file we just created:

```
alpha_rarefaction.py -i otus/otu_table.biom -m Pasting_Map.txt -o wf_arareo/ -p alpha_params.txt -t otus_rep_set.tre
```

Descriptions of the steps involved in alpha_rarefaction.py follow:
5. Yabish
5.2 Resulting job in Front End
Session Two – Administering Yabi (60 mins)
- Workflow management
- Resources
- Users and Credentials
- Tools
1. Workflow Summaries
1. “Syslog”

Select syslog to change

2.1 Backends

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Scheme</th>
<th>Hashname</th>
<th>Port</th>
<th>Path</th>
<th>Uri</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3 fs</td>
<td>S3 filesystem</td>
<td>s3</td>
<td>yabi-sing.amazonaws.com</td>
<td></td>
<td></td>
<td>s3://yabi-sing.amazonaws.com/</td>
<td>View</td>
</tr>
<tr>
<td>EPIC fs</td>
<td>EPIC sftp filesystem</td>
<td>sftp</td>
<td>epic leve.org</td>
<td>22</td>
<td>/</td>
<td>sftp://epic leve.org:22/</td>
<td>View</td>
</tr>
<tr>
<td>EPIC ev</td>
<td>EPIC ssh execution</td>
<td>ssh</td>
<td>epic leve.org</td>
<td>22</td>
<td>/</td>
<td>ssh://epic leve.org:22/</td>
<td>View</td>
</tr>
<tr>
<td>Local Execution</td>
<td>This backend gives access to execution on the machine running Yabi.</td>
<td></td>
<td>local hosted</td>
<td></td>
<td></td>
<td></td>
<td>View</td>
</tr>
<tr>
<td>Local Filesystem</td>
<td>This backend gives access to the file system on the machine running Yabi.</td>
<td></td>
<td>local hosted</td>
<td></td>
<td></td>
<td></td>
<td>View</td>
</tr>
<tr>
<td>null backend</td>
<td>Use this backend when tools should not be run i.e fileselector</td>
<td>null</td>
<td>localhost.localdomain</td>
<td></td>
<td></td>
<td>localhost.localdomain:</td>
<td>View</td>
</tr>
</tbody>
</table>

8 backends
### Change backend

<table>
<thead>
<tr>
<th>Name</th>
<th>EPIC ex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>EPIC ssh+pbspro execution</td>
</tr>
<tr>
<td>Scheme</td>
<td>ssh+pbspro</td>
</tr>
<tr>
<td>Hostname</td>
<td>epic.iVEC.org</td>
</tr>
<tr>
<td>Port</td>
<td>22</td>
</tr>
<tr>
<td>Path</td>
<td>/</td>
</tr>
</tbody>
</table>

**Change backend description:**

- **Scheme:** ssh+pbspro
  - Must be one of global, sge, torque, ssh, ssh+pbspro, ssh+torque, ssh+sge, localsh, exclusive, null, http, https, gridftp, sftp, scp, s3, localfs, null.
- **Hostname:** epic.iVEC.org
  - Hostname must not end with a ‘/’.
- **Port:** 22
- **Path:** /?
  - Path must start and end with a ‘/’.
  - Execution backends must only have / in the path field.
  - For filesystem backends, Yubi will take the value in path and combine it with any path unique in Backend Credential to form a URL. I.e. http://remserver.<domain>/home would be entered here and then on the Backend Credential for User X you would enter their home directory in the User Directory field i.e. UserX/. This would then combine to form a valid URL: http://remserver.<domain>/home/UserX/
- **Max connections:**
  - Backend connection limit. Does not affect front and immediate mode requests. Blank means no limit on the number of connections. ‘0’ means no connections allowed (frozen).

- **Yubikey supported:**
  - Backend supports ‘ip’ localised copies.
- **Link supported:**
  - Backend supports ‘ln’ localised (symlinking).

**Submission:**

- `#!/bin/sh`
- `PBS = -l walltime=`
- `PBS = -l memory=`
- `PBS = -l nodes=`
- `PBS = -W group_list=`
- `# for module or modules`
- `module load [module]`
- `# for module in modules`
- `ls -d /`
- `ls -d /##Working/`

**Submission note:**

Make script to be used to generate the submission script. (Variables: walltime, memory, cpus, working, modules, command)
2.2.1 Credentials

Change credential

<table>
<thead>
<tr>
<th>Description</th>
<th>starcluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Username</td>
<td>ubuntu</td>
</tr>
<tr>
<td>Password</td>
<td>$AESS$</td>
</tr>
<tr>
<td>Cert</td>
<td>$AESS$de6342cf57f080e6a213ecbcb9baac35</td>
</tr>
<tr>
<td>Key</td>
<td>$AESS$b627ce8b6e7be85c35460f205189bba590b1730ba33a6445582eb36e2ed7c5045bf254bf5b334547e53658f82c62f76f4643b3d0b2d76c6b0f0b2b76f0d86f61fde4d4d0d19ab7734c2db6dc63e61e976379b128d5b4d0d81346c1301547bf72d0d6d3bf84f0d07e81d0f2e07b07326b841d8e76e29d29f7f7c3094192b2e4f66da819f4f726b4e0b1d40556341c10d2f43761e5557ce06e0b8a132761347bacc9f13a55504349631edcc6d6e7fedce2653818717650e44c275046a26e9296b7f917680c31c0d892e6324cc8e0b0365a594f850c942a78aeb361331843a4cc7c8d74a595b042762139a87960ef440c7052e57fbb1edc1162248762a23096e19467f44e22790254109c75146370e6b4176b732512f11054f6bda9267d433bb919f2c044e8b49ebe0857eb2247ee46e7736f3c935a06261af5290abb9eb5c4e8edd6f2998e8ed8e6f1cefa68442b80fca0a423510df8a880ca</td>
</tr>
</tbody>
</table>

User: andrew

Expires on: 2112-01-25

Yubi username.

Delete
2.3 Backend Credential

Change backend credential

Backend:  
Creden,al

Credential:

User Directory:

Visible

Default stageout

Submission:

Delete

11/20/12
## 2.4 Adding Tools

### Change tool

<table>
<thead>
<tr>
<th>Name</th>
<th>split_libraries.py</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique tool name for internal use.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Display name</th>
<th>split_libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool name visible to users.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Path</th>
<th>split_libraries.py</th>
</tr>
</thead>
<tbody>
<tr>
<td>The path to the binary for this file. Will normally just be binary name.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description</th>
<th>Split libraries according to barcodes specified in mapping file</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The description that will be sent to the frontend for the user.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enabled</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable tool in frontend.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The execution backend for this tool.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fs backend</th>
<th>starcluster fs - scp://sing-yabi-sc-prod-19-09-2012-master.sc.ccs.qs.unsw.edu.au:22/yabidata/</th>
</tr>
</thead>
<tbody>
<tr>
<td>The filesystem backend for this tool.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accepts input</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>If checked, this tool will accept inputs from prior tools rather than presenting file select widgets.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cpus</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Walltime</th>
<th>01:00:00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Module</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 2.4 Tool Parameters

<table>
<thead>
<tr>
<th>Tool Parameters</th>
<th>Tool Parameter: --map</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Switch</strong>:</td>
<td>--map</td>
</tr>
<tr>
<td></td>
<td>The actual command line switch that should be passed to the tool i.e. -i or --input-file</td>
</tr>
<tr>
<td><strong>Switch use</strong>:</td>
<td>both</td>
</tr>
<tr>
<td></td>
<td>The way the switch should be combined with the value.</td>
</tr>
<tr>
<td><strong>Rank</strong>:</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>The order in which the switches should appear. Leave blank if order is unimportant.</td>
</tr>
<tr>
<td><strong>Mandatory</strong></td>
<td>Select if the switch is required as input.</td>
</tr>
<tr>
<td><strong>Hidden</strong></td>
<td>Select if the switch should be hidden from users in the frontend.</td>
</tr>
<tr>
<td><strong>Output file</strong></td>
<td>Select if the switch is specifying an output file.</td>
</tr>
<tr>
<td><strong>Extension param</strong>:</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>If an extension is selected then this extension will be appended to the filename. This should only be set for specifying output files.</td>
</tr>
<tr>
<td><strong>Possible values</strong>:</td>
<td></td>
</tr>
</tbody>
</table>
### 2.4.2 Tool Groupings

<table>
<thead>
<tr>
<th>Name</th>
<th>Tools in toolgroup, by toolset</th>
</tr>
</thead>
<tbody>
<tr>
<td>qiime</td>
<td>qiime: (alpha_rarefaction.py, beta_diversity.py, beta_diversity_through_pi) demo: (alpha_rarefaction.py, beta_diversity.py, beta_diversity_through_pi)</td>
</tr>
<tr>
<td>bioven</td>
<td>bioven: (ccg_pi_db_import.py, msblast, cgg_pihmmlparser.py)</td>
</tr>
<tr>
<td>genomics</td>
<td>alltools: (chips)</td>
</tr>
<tr>
<td>ebi services</td>
<td>alltools: (ebi_ncbiblast, ebi_clustalw2, ebi_fasta, ebi_wublast, ebi_fasta_2, ebi_wublast_2, ebi_fasta_fasta_2) demo: (ebi_ncbiblast, ebi_clustalw2, ebi_fasta, ebi_wublast, ebi_fasta_2)</td>
</tr>
<tr>
<td>emboss</td>
<td>alltools: (extractseq, getorf, transeq, eprimer3) demo: (eprimer3)</td>
</tr>
<tr>
<td>system</td>
<td>alltools: (hostname, hostname epic) demo: (hostname)</td>
</tr>
<tr>
<td>blast</td>
<td>alltools: (blast (epic), blast (starchuster))</td>
</tr>
<tr>
<td>select data</td>
<td>alltools: (fileselector) demo: (fileselector)</td>
</tr>
</tbody>
</table>

8 tool groups
2.4.2 Tool Sets

Change tool set

Name: demo

Users:

Available users

admin
bioven

Chosen users

andrew
demo
demo01
demo10
demo11
demo12
demo13
demo14
demo15
demo16
demo17
demo18
demo19

Choose all
Clear all

Delete
2.4.3 Tool Export

```python
( "tool": {
  "outputExtensions": [
    { "must_be_greater_than": false, "must_exist": false, "file_extension_patterns": "*.txt" },
    { "must_be_greater_than": false, "must_exist": false, "file_extension_patterns": "*.faa" }
  ],
  "display_name": "split Libraries",
  "name": "split_library.py",
  "parameter_list": []
  { "default_value": ",", "mandatory": true, 
    "use_output_filename_switch": false, 
    "output files": false, 
    "possible_values": null, 
    "task": 9,
    "switch_use_description": "both the switch and the value will be passed in the argument list. They will be separated by a space.";
    "switch_use_file_assignment": "file specifies the headerstring",
    "switch": "--map",
    "acceptedExtensionList": [ "", "" ]
  },
  { "default_value": ",", "mandatory": true,
    "use_output_filename_switch": false, 
    "output files": false, 
    "possible_values": null, 
    "task": 9,
    "switch_use_description": "both the switch and the value will be passed in the argument list. They will be separated by a space.";
    "switch_use_file_assignment": "file specifies the headerstring",
    "switch": "--map",
    "acceptedExtensionList": [ "", "" ]
  }
],
"help_text": "Mapping file. NOTE: Must contain a header line indicating SampleID in the first column and BarcodeSequence in the second, Linker Primer Sequence in the third.",
"batch_bundle_files": false,
"file_assignment": "batch",
"hidden": false,
"id": 82,
"switch_use_display_text": "both"

open_secondary: "";
"accept_secondary": false;
"files": [ "", "", "", "", ""]
}
"help_text": "Fastq files.",
"batch_bundle_files": false,
"file_assignment": "batch",
"hidden": false,
"id": 82,
"switch_use_display_text": "both",
"accept_secondary": false;
"files": [ "", "", "", "", ""]
}
```

11/20/12
2.4.4 Parameter Switch Use

<table>
<thead>
<tr>
<th>Parameter Switch Use</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>mv</code> <code>-i</code>, <code>-input/switch</code></td>
<td>Use this when you need to move files from input directory to output directory, typically when scripts or tools output where the input files are.</td>
</tr>
<tr>
<td><code>cp</code> <code>-i</code>, <code>-input/switch</code></td>
<td>Use this when you need to copy files from input directory to output directory, typically when scripts or tools output where the input files are.</td>
</tr>
<tr>
<td><code>&gt;</code>, <code>&gt;</code> <code>-value(s)</code></td>
<td>Use this to redirect the output of <code>stdout</code> into a file.</td>
</tr>
<tr>
<td><code>%switch(s)=%value(s)</code></td>
<td>Both the switch and the value will be passed in the argument list. They will be separated joined with an equals(=) character with no spaces.</td>
</tr>
<tr>
<td><code>pair</code></td>
<td>The switch and the value passed in to the argument list as a pair.</td>
</tr>
<tr>
<td><code>nothing</code></td>
<td>The switch and the value won't be passed in the argument list.</td>
</tr>
<tr>
<td><code>%switch(s)=%value(s)</code></td>
<td>Both the switch and the value will be passed in the argument list. They will be joined together with no space between them.</td>
</tr>
<tr>
<td><code>%switch(s)=%value(s)</code></td>
<td>Both the switch and the value will be passed in the argument list. They will be separated by a space.</td>
</tr>
<tr>
<td><code>valueOnly</code></td>
<td>Only the value will be passed in the argument list (i.e., the switch won't be used).</td>
</tr>
<tr>
<td><code>switchOnly</code></td>
<td>Only the switch will be passed in the argument list.</td>
</tr>
</tbody>
</table>
2.5 Automatically add tools

class ToolBuilder(EBI):

    """This class can be used from the python interpreter to create new yabi tool json.

    from ebi.ebi_rest import EBI, ToolBuilder, start_logging
    start_logging()
    tb = ToolBuilder(uri='http://www.ebi.ac.uk/Tools/services/rest/wublast/', name="ebi_wublast", display_name="Wublast")
    tb.build()
    tb.write_json()
    """

def __init__(self, *args, **kwargs):

11/20/12
Overview cont...

Session Three – Installing Yabi (60 mins)
- Local installation for testing
- Deploying to the cloud
Open Source

A sophisticated online research environment for Grid, High Performance and Cloud computing.

Workflows made easy
A sophisticated online research environment for Grid, High Performance and Cloud computing.

About
A fast application stack to provide users with an intuitive, easy-to-use abstraction of compute and data environments. Developed at the Centre for Comparative Genomics. Yabi has been deployed across a diverse set of scientific disciplines and high performance computing environments.

In a hurry? Try our Quickstart Page. Otherwise take a look at the documentation.

Citation

Licence
GNU GPL v3 for non-commercial use only. Please contact the Centre for Comparative Genomics if you require a licence other than GPL for legal or commercial reasons.

Contact Us
Yabi Quickstart

```
Terminal: andrew@ubuntu: ~/hg-devel/yabi-quickstart

andrew@ubuntu: ~/hg-devel/yabi-quickstart$ source virt_yabi-quickstart/bin/activate
(yabi-quickstart)andrew@ubuntu: ~/hg-devel/yabi-quickstart$ fab -l
Available commands:

admin_activate_config
admin_active_config
admin_bootstrap
admin_createdb
admin_deactivate_config
admin_dropdb
admin_initdb
admin_jslint
admin_kllcelery
admin_kllserver
admin_list_configs
admin_quickstart
admin_recreatedb
admin_require
admin_runcelery
admin_runserver
admin_select_test_config
admin_selected_test_config
admin_tests
be_bootstrap
be_createdirs
be_kllserver
be_quickstart
be_runserver
clean
dllservers
quickstart
runserver
runtests

```

Activate the passed in Yabi admin config
Displays the active Yabi Admin config
Bootstrap the yabadmin project
Create the DB of Yabadmin
Deactivates the active Yabi Admin config
Drop the DB of Yabadmin
Initialise the DB of Yabadmin
Runs Google Closure Linter on JavaScript in ...
Kill the yabadmin celery server
Kill the yabadmin local server
Displays available configs for Yabi Admin
Quickstart the yabadmin project (bootstrap,...
Recreate (drop then create) the DB of Yablad...
Install additional requirements into the yab...
Run the yabadmin celery server for local dev ...
Run the yabadmin server for local dev (db ... Select the passed in config to be used when ...
Displays the configuration used for running ...
Runs all the tests in the yabadmin project
Bootstrap the yabbe project
Creates necessary directories for the yabbe...
Kill the yabbe local server
Quickstart the yabbe project (bootstrap, cr ...
Run the yabbe server for local dev (db for ...
Clean all virtual environment directories
Kills all the local development servers in t...
Quickstart the whole YABI stack (admin, be, ...
Run all servers in the YABI stack for local ...
Run all the YABI tests

11/20/12