

Florida Atlantic Harbor Branch Oceanographic Institute  
Attn. Mrs. S. Edge  
P.O. Box 5600  
NORTH FORT PIERCE Florida FL 34946  
VER. ST. VAN AMERIKA

Dear Mrs. Edge,

**Subject**

Project 54114

**Your Reference**

**Date**

6 February 2013

**Contact**

Walter Pirovano

By way of this letter we would like to inform you that the data for Project 54114 has been uploaded onto our sFTP server. Below you will find all relevant information to retrieve the files. In addition a project report is provided in the attachment of the email.

We would like to thank you for choosing BaseClear. We truly hope that the results meet your expectations and that you will consider us also for future projects.

If you have any further questions or need further assistance please do not hesitate to contact me.

With kind regards,



Walter Pirovano  
Product specialist bioinformatics and genome analysis

✉ [walter.pirovano@baseclear.com](mailto:walter.pirovano@baseclear.com)

**Reference**

Project 54114

**Page**

2 / 3

**Samples**

The project data corresponds to the following samples:

- Group 1
  - 25ppt
  - 30ppt
  - 35ppt
- Group 2
  - Buc
  - Bod
  - Hoc

**Project summary**

The goal of this project was to perform a Next Generation sequencing project including the following services:

- Library preparation
- Next Generation sequencing
- Bioinformatics analysis including:
  - Quality analysis of FASTQ sequence reads
  - *De novo* transcriptome assembly and annotation
  - RNA-Seq analysis

**Authorisation**

Walter Pirovano  
(Product specialist)

**Project report**

A Next generation sequencing project report can be found in the attachment of the email.

**Accessing the sFTP server**

The sFTP server provides a safe option for data download as connections are encrypted through a secure shell. You can connect to it using an sftp-client. Our preference is FileZilla which can be downloaded from <http://filezilla-project.org/download.php?type=client>.

After installation, open the program and login using the following account details:

- host / address: sftp.baseclear.com
- username: 54114
- password: LiucIth4
- port: 22

Once you are connected, you can drag the files from the server (on the right) to your own environment (on the left).

The data will be available on our sFTP server for the coming two weeks.

**Reference**

Project 54114

**Page**

3 / 3

**Folder structure on the server**

- **raw\_sequences**  
Contains the zipped sequence reads in FASTQ format after analysis with CASAVA and a post-filtering script (to remove low quality, PhiX-control and adapter reads). These can be unzipped using either WinZip ([www.winzip.com](http://www.winzip.com)) or 7-zip ([www.7-zip.org](http://www.7-zip.org)) if the files are > 6 GB.
- **transcriptome\_assembly**  
Contains the de novo assembly(ies) made with Trinity transcriptome assembly software. For each group, larvae/salinity (-> 25ppt, 30ppt, 35ppt) and adult coral (-> buc, bod, hoc). The contig sequences are provided in FASTA format.  
In addition for each assembly the corresponding annotation file in GFF format is provided. Annotations are made using BLAST against the UniProt/SwissProt database.
- **reference\_alignment**  
Contains the RNA-Seq reference alignment(s) made with the CLC Genomics Workbench. The alignments are provided in BAM format.
- **RNA-seq\_analysis**  
Contains the RNA-Seq expression analysis files made with the CLC Genomics Workbench. Files include per sample the expression table in Excel format. Also a summarizing analysis table is provided containing comparative expression values and statistical analysis between the samples.
- **analysis\_figures**  
Contains the quality control analysis figures of the RNA-Seq experiment. These include the RPKM expression comparisons, Principle Component Analysis, Hierarchical clustering plots in PNG format.