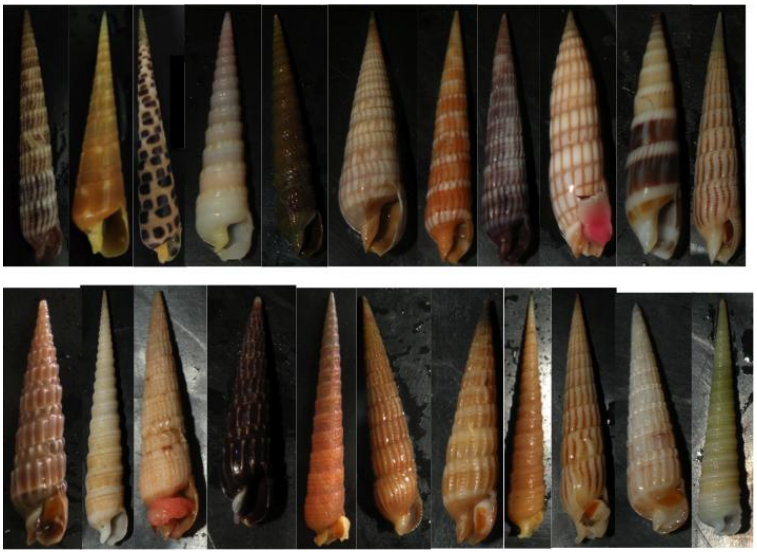


Plate 1: Representative terebrid species collected in Inhaca

Top panel: *Myurella kilburni, Acus dimidiata, Terebra subulata, Terebra quoygaimardi(W), Cinguloterebra anilis, Perirhoe cerithina, Myurella columellaris, Myurella nathaliae, Hastula solida, Myurella undulata,*

Bottom panel: *Striterebrum calignosa, Terebra funiculata, Hastulopus conspersa, Duplicaira mozambiquensis, Terebra punctostriata, Terebra succincta, Myurella joserosadoi, Terebra laevigata, Myurella undulata, Duplicaria baileyi,Cinguloterebra hoaraui*

Specific Aim #2. Identify transcripts coding for terebrid neuropeptides and organize terebrid peptides into gene superfamilies

Construction of cDNA libraries via cloning and Sanger sequencing has been the method of choice for analyzing the transcriptomes of non-model organisms. However, recent advances in the commercially available deep RNA sequencing (RNA-Seq) technologies of Illumina and Roche have made it possible to expand the level of sequence coverage while solving the problems of de novo assembly[10, 11]. Illumina technology currently provides 250 base pair reads, shorter than those of the Roche platform, but superior in a number of other areas such as number of reads, accuracy in base calling and the ability to sequence homopolymeric regions. Advances in the software programs available to perform contig assembly and blasting of large datasets, as well as the use of tailored scripts to mine data for specific types of transcripts (i.e. cysteine rich) have made it feasible to obtain substantial data both on peptide toxin transcripts and the expression of genes involved in myriad cellular processes. These deep sequencing technologies have already been performed on cone snail species such as *C. bullatus*, *C. consors*, and *C. geographus* [12-14]. While there are still difficulties inherent in de novo assembly, the discovery of new peptide toxin transcripts and an understanding of other proteins expressed in the venom duct are growing rapidly as a result of these efforts.

Given the recent RNA-Seq results with cone snails, a next generation, high-throughput approach was undertaken to perform sequencing and de novo transcriptome assembly of *Terebra anilis* venom ducts dissected from specimens collected during the fall 2011 trip to Inhaca, Mozambique. We chose the Illumina platform for this sequencing project given the reasons cited above. Total RNA was extracted from eight *T. anilis* venom ducts with Qiagen RNeasy kit and 10ng of this product was used with Clontech's SMARTer™ Ultra Low RNA Kit for

Illumina Sequencing to perform 1st strand cDNA synthesis and amplification. *T. anilis* cDNA was prepared with the Kapa Biosystems Kit (fragmentation and adapter ligation) for sequencing on two paths of an 8-lane Illumina flow cell. Throughout the process, total RNA, starting cDNA and cDNA library generation were subjected to quantification and quality control on an Agilent 2100 Bioanalyzer (Fig. 4).

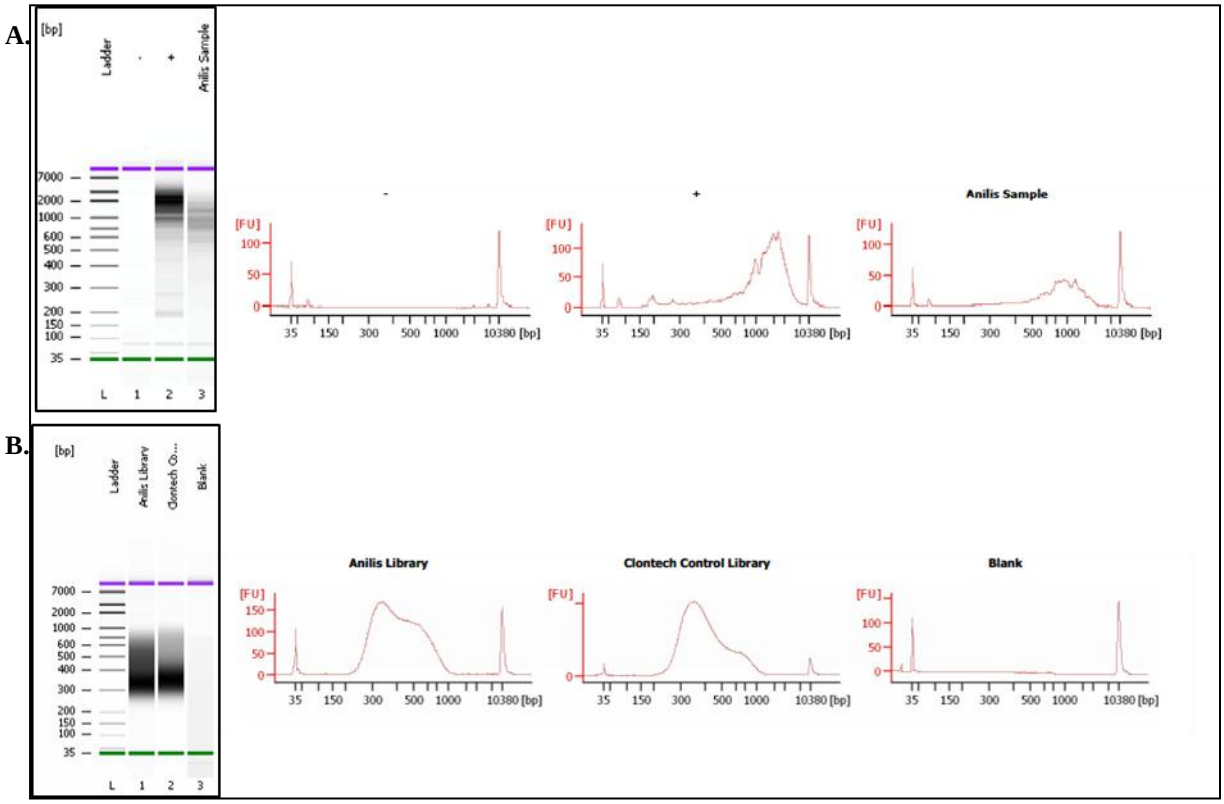


Figure 4. Agilent Bioanlyzer trace of *T. anilis* cDNA , control mouse cDNA, and negative control after PCR amplification of 1st strand cDNA with SMARTer™ Ultra Low RNA Kit (A). Bioanalyzer trace after cDNA library prep

Illumina sequencing of *T. anilis* venom duct cDNA yielded 288,959,674 reads of 100 bp length from lane 6, and 248,093,494 reads from lane 7 (Table 1.).

Table 1. Statistical summary of 100 base pair Illumina raw reads of *Terebra anilis* venom ducts.

Sample	Organism/Source	Application	Submission(s)	Demultiplexing Options	# Reads	% bases ≥ Q30	Mean QS (PF)
Anilis Total RNA (extracted 7-25-2012)	C. anilis	RNA-Seq	C0N4WACXX (Lane 6)	(mismatch: 1) (specific_barcodes)	288,959,674	89.43%	34.64
			C0N4WACXX (Lane 7)	(mismatch: 1) (specific_barcodes)	248,093,494	89.43%	34.64

Subsequent contig assembly of raw reads from lane 6 was performed using Trans-ABySS 1.4.4 to generate 411,826 contigs > 250 bp and 38,232 contigs > 100 bp, with an N50 length of 326 bp. This dataset was analyzed first by blasting against a comprehensive set of conotoxins downloaded from Conoserver and then

RNA extraction for fabulous results – LW – Holford Lab, Hunter College

General notes: Change gloves obsessively. Work in a hood. Use RNase AWAY® and DNA AWAY™ to wipe down surfaces. UV crosslink pipettes, tips etc.

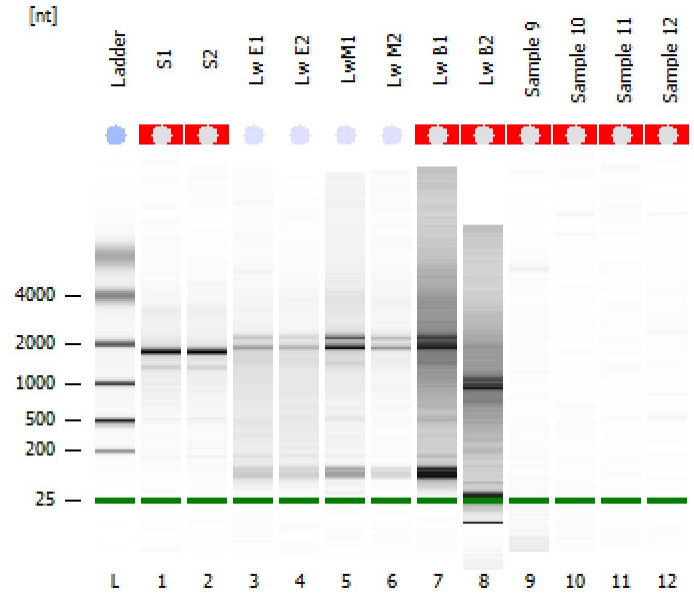
Protocols for Trizol + PureLink RNA Mini Kit
http://tools.invitrogen.com/content/sfs/manuals/Trizol_Plus_man.pdf

1. Pulverize tissue in 1.7ml eppendorf tube. Use long pestles to grind tissue up, keep tissue cold by immersing bottom part of tube in liquid N2. As soon as tissue is suitably smashed up (powder) add Trizol. The amount of Trizol added will depend on the amount of starting tissue
   NB: I put tissue directly in Trizol and use FischerSci PowerGen 125 to homogenize tissue (Qiagen TissueRuptor also works). I think this works best!
2. Add 1ml (large sample) or 500 µl (small sample) of Trizol. Vortex to mix and let sit at bench for a couple of minutes (*see protocol for appropriate ratio of mg tissue/Trizol and correct reagent amounts from this step on down*)
3. Add 200 µl (if starting with 1ml of Trizol), or 100 µl (if used 500 µl of Trizol) of chloroform. Invert for about 20 seconds, let sit at RT for 3 minutes
4. Spin at 12K G in 4 deg centrifuge for 15 minutes
5. Collect top colorless layer (containing RNA) be sure not to get any solution from bottom layer or solid phase. Pipette into a new eppe.
   **optional* steps 6-8 (can provide better separation)*
6. Add 500 µl (if starting with 1ml of Trizol) or 250 µl (if starting with 500 µl of Trizol) of isoproponyl. Mix by inversion and let sit on lab bench for 10 min.
7. Spin at 12K G for 10 mins at 4 deg
8. Remove supernatant being careful not to disturb the pellet
9. Add equal volume of 70% ETOH, and mix well by vortexing
10. Invert tube to resuspend any precipitate and proceed to PureLink RNA Mini Kit Protocol (and make sure to follow DNase treatment listed as optional)
11. Run purified total RNA on Agilent Bioanalyzer to determine quality and concentration.
12. Store RNA -80!!!

Assay Class: Eukaryote Total RNA Nano
Data Path: C:\...Eukaryote Total RNA Nano_DE13804219_2013-04-19_12-18-57.xad

Created: 4/19/2013 12:18:57 PM
Modified: 4/19/2013 12:42:50 PM

Electrophoresis File Run Summary



Instrument Information:

Instrument Name: DE13804219 Firmware: C.01.069
Serial#: DE13804219 Type: G2939A

Assay Information:

Assay Origin Path: C:\Program Files (x86)\Agilent\2100 bioanalyzer\2100 expert\assays\RNA\Eukaryote Total RNA Nano Series II.xsy

Assay Class: Eukaryote Total RNA Nano

Version: 2.6

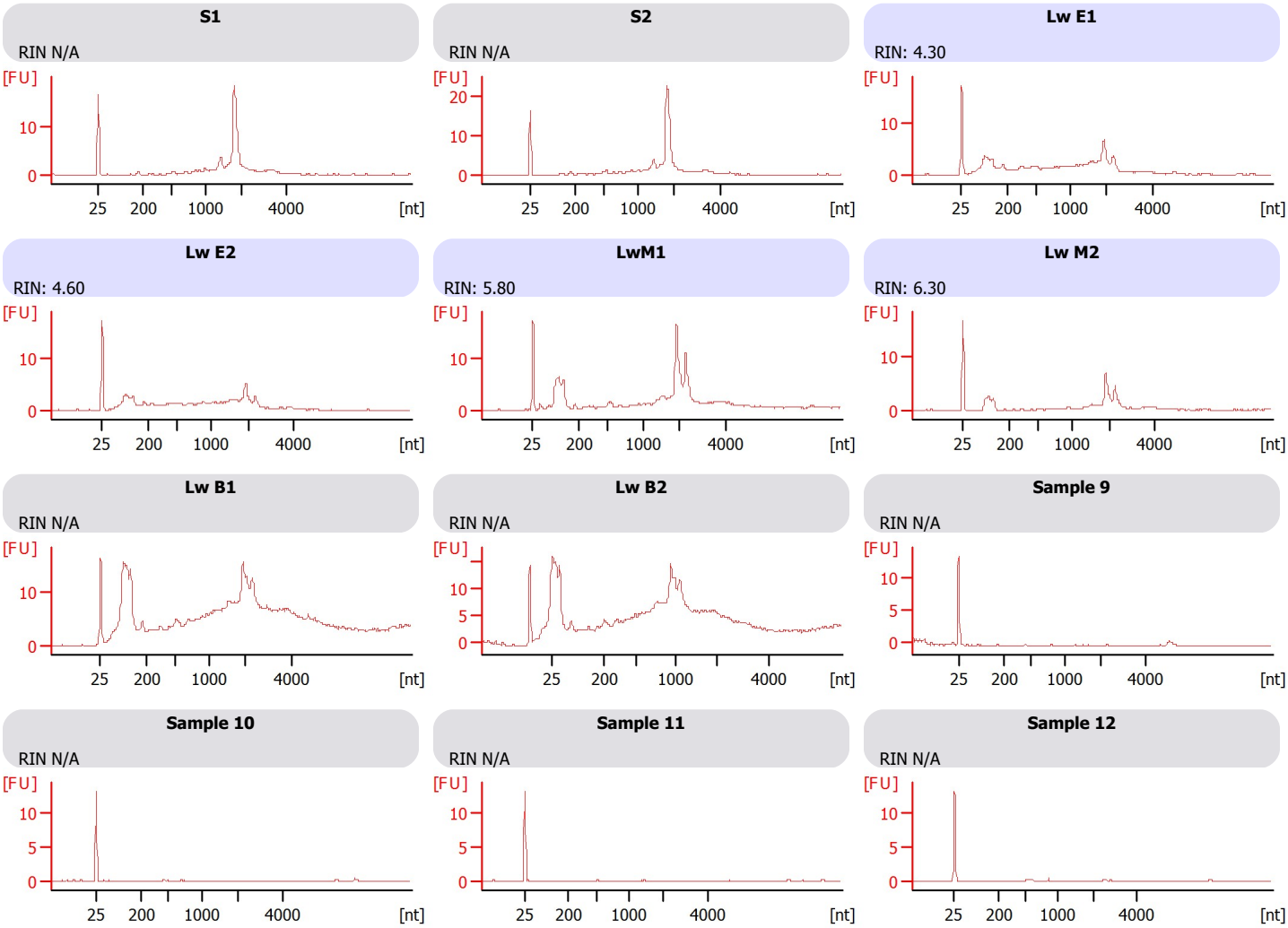
Assay Comments: Total RNA Analysis ng sensitivity (Eukaryote)

Chip Information:

Chip Lot #:

Reagent Kit Lot #:

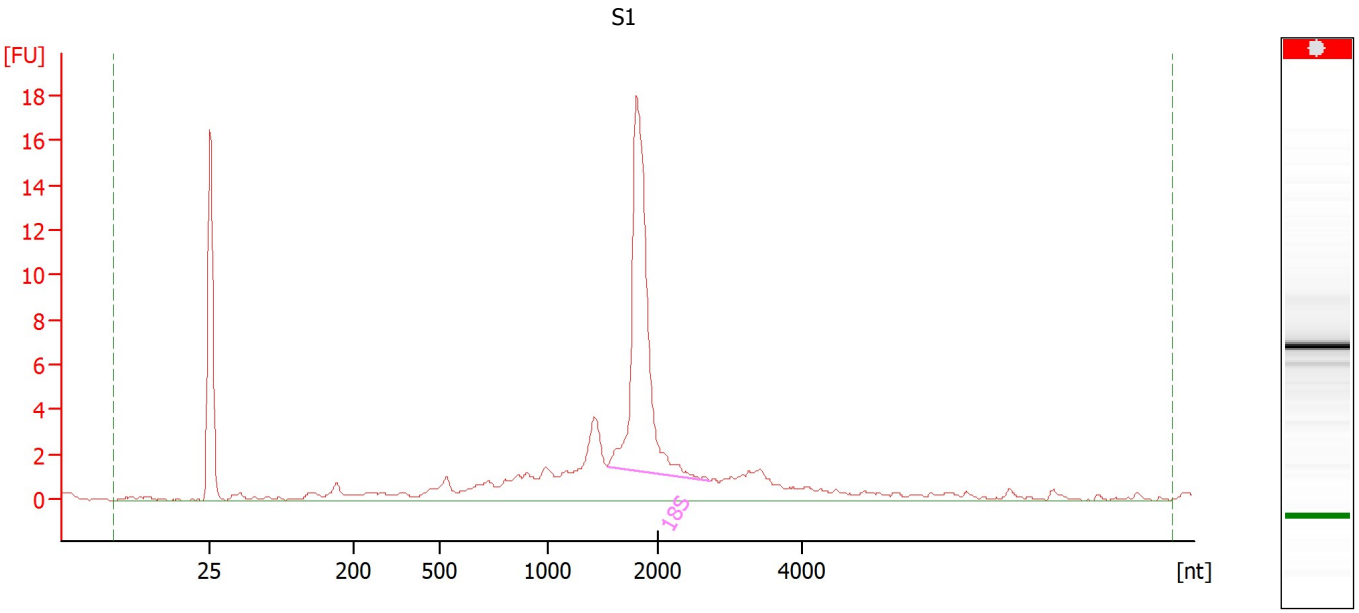
Chip Comments:



Assay Class: Eukaryote Total RNA Nano
Data Path: C:\...Eukaryote Total RNA Nano_DE13804219_2013-04-19_12-18-57.xad

Created: 4/19/2013 12:18:57 PM
Modified: 4/19/2013 12:42:50 PM

Electropherogram Summary Continued ...



Overall Results for sample 1 : S1

RNA Area:	90.4	RNA Integrity Number (RIN):	N/A (B.02.08)
RNA Concentration:	30 ng/µl	Result Flagging Color:	
rRNA Ratio [28s / 18s]:	0.0	Result Flagging Label:	RIN N/A

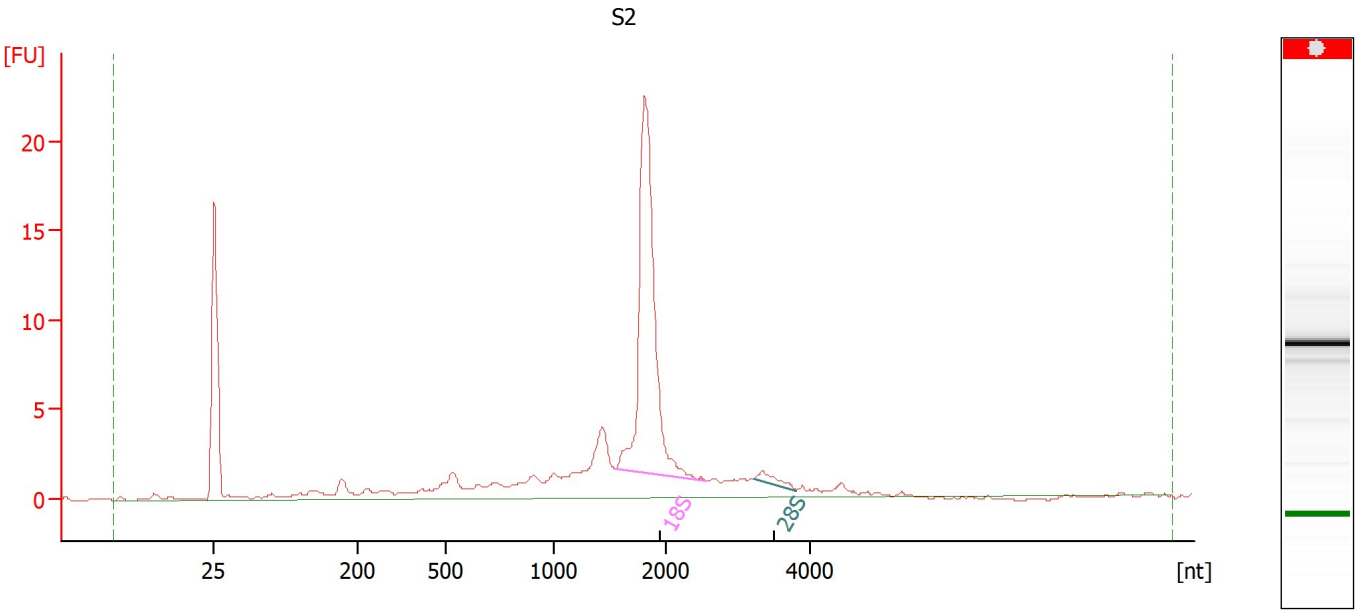
Fragment table for sample 1 : S1

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
18S	1,542	2,707	31.2	34.5

Assay Class: Eukaryote Total RNA Nano
Data Path: C:\...Eukaryote Total RNA Nano_DE13804219_2013-04-19_12-18-57.xad

Created: 4/19/2013 12:18:57 PM
Modified: 4/19/2013 12:42:50 PM

Electropherogram Summary Continued ...



Overall Results for sample 2 : S2

RNA Area:	99.1	RNA Integrity Number (RIN):	N/A (B.02.08)
RNA Concentration:	33 ng/µl	Result Flagging Color:	
rRNA Ratio [28s / 18s]:	0.0	Result Flagging Label:	RIN N/A

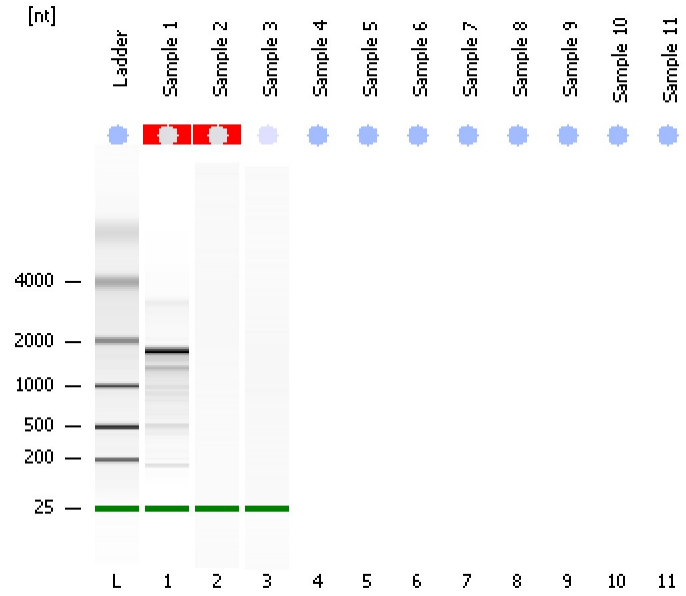
Fragment table for sample 2 : S2

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
18S	1,551	2,545	38.1	38.5
28S	3,199	3,808	1.3	1.3

Assay Class: Eukaryote Total RNA Pico
Data Path: C:\...Eukaryote Total RNA Pico_DE54700454_2012-07-26_13-14-45.xad

Created: 7/26/2012 1:14:45 PM
Modified: 7/26/2012 1:26:27 PM

Electrophoresis File Run Summary



Instrument Information:

Instrument Name: NYUGAC
Firmware: C.01.069
Serial#: DE54700454
Type: G2939A

Assay Information:

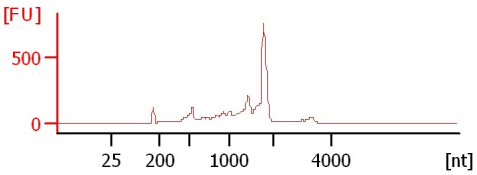
Assay Origin Path: C:\Program Files\Agilent\2100 bioanalyzer\2100 expert\assays\RNA\Eukaryote Total RNA Pico Series II.xsy
Assay Class: Eukaryote Total RNA Pico
Version: 2.6
Assay Comments: Total RNA Analysis pg sensitivity (Eukaryote)

Chip Information:

Chip Lot #:
Reagent Kit Lot #:
Chip Comments:

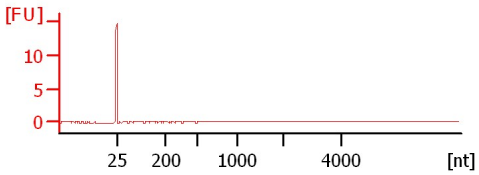
Sample 1

RIN N/A



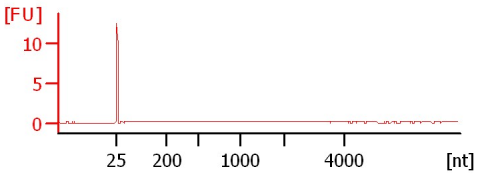
Sample 2

RIN N/A



Sample 3

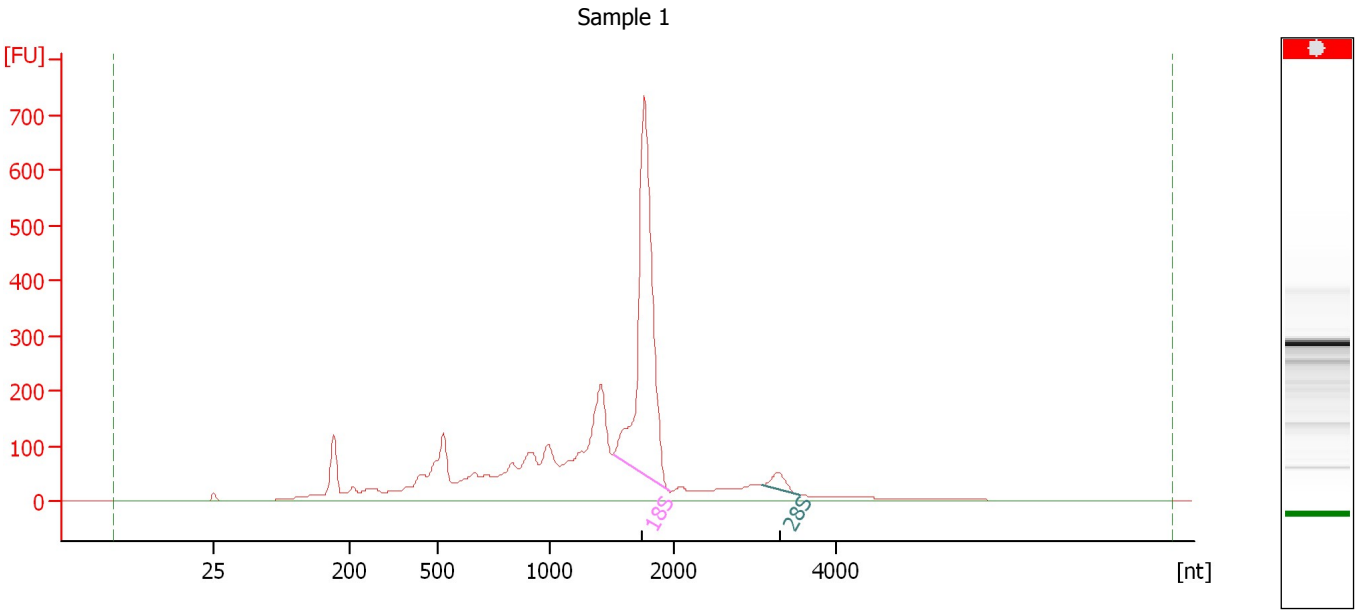
RIN:1



Assay Class: Eukaryote Total RNA Pico
Data Path: C:\...Eukaryote Total RNA Pico_DE54700454_2012-07-26_13-14-45.xad

Created: 7/26/2012 1:14:45 PM
Modified: 7/26/2012 1:26:27 PM

Electropherogram Summary Continued ...



Overall Results for sample 1 : Sample 1

RNA Area:	4,081.6	RNA Integrity Number (RIN):	N/A (B.02.08)
RNA Concentration:	16,407 pg/µl	Result Flagging Color:	
rRNA Ratio [28s / 18s]:	0.0	Result Flagging Label:	RIN N/A

Fragment table for sample 1 : Sample 1

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
18S	1,505	1,963	1,148.2	28.1
28S	3,081	3,558	47.1	1.2