Preparation of Illumina libraries from pMC-MERII-6xPROM-luc-BC **transfected cells**

Step 1.0 RNA extraction

RNA is extracted with TriZol and purified with LiCl precipitation.

1. Collect U2OS cells (3.5cm dish) in 1 mL Trizol. store at -80°C until RNA extraction.
2. Extract RNA according to trizol protocol.
3. Resuspend the RNA in 40 uL water.
4. Add 20 uL of LiCl (LiCl 7.5M, EDTA 50mM, autoclaved).
5. Mix well and precipitate at -20°C o/n.
6. Centrifuge max speed at 4°C for 30 minutes.
7. Wash the RNA pellet with ice-cold EtOH 75%.
8. Centrifuge max speed at 4°C for 15 minutes.
9. Dry the pellet and resuspend in 30 uL water. Measure the concentration at nanodrop.
10. The RNA is stored at -80 °C until the reverse transcription.

Step 1.1 Reverse Transcription with Oligo(dT) tail

1. Thaw LiCl-precipitated **U2OS RNA** (~300 ng/uL) and the reactives on ice. (keep SuperScript and RNAseOUT at -20°C).
2. Prepare **2 ug** of RNA + 1 uL **Primer 81** (NheI-Oligo(dT)-VN, 10 uM) in 0.2ml tubes and complete with H2O to a final volume of **12 ul**.
3. Put the RNA tubes in thermocycler and start RT program (65°C for 10', 4°C hold), to denature RNA secondary structure.
4. While the thermocycler is running, prepare the RT mastermix (**MM**):

Number of reactions		10
	uL/reaction	uL for RT mix
dNTP mix (25 mM)	0.42	4.2
First-strand buffer 5 x	4.2	42.0
DTT 0.1 M	2.1	21.0
Water	0.44	4.4
RNaseOUT	0.21	2.1
Superscript II	1.05	10.5

5. Recover tubes from thermocycler
6. Distribute **8 uL MM** to **tubes**
7. Plan **no-RT** reaction controls as appropriate.
8. Return to thermocycler and continue the program (press Enter, anneal 25°C for 10 min, elongate 37°C for 20 min and 42°C for 30 min, inactivate at 65°C for 10 min, hold to 4°C).
9. The cDNA is stored at -20 °C until the RT-PCR.

Step 1.2 – **qPCR1**: Inserting the left Illumina end (Index)

With this first PCR we amplify a luciferase fragment containing the barcode. The forward primer is specific for each sample and introduces an index sequence to recognize the sample inside the illumine reads. The reverse primer anneals to the Oligo(dT) tail, making the amplification specific for cDNA. The plasmid DNA cannot amplify (as controlled in a no-RT reaction). To maximize library coverage, 1 ug cDNA is amplified in 4 separate reactions of 250 ng cDNA template each. The PCR is repeated twice (2 ug total).

Expected fragment 266 bp:

GACGCTCTTCCGATCTNNNNNN**CAAGATCGCCGTGTAATAAT**bvhdhhhhbbvvddhhhhddTCTAGAGTC
GGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAA
TGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGC**A**
ATAAACAAGTTAACAACAAAAAAAAAAAAAA**CGTGCTAGCCGTGCGGCCGCTCA**

1. Thaw the cDNA tubes and the reagents on ice
2. Add 21 uL Tris 10 mM pH 8.0 to each tube, mix well.
3. Distribute 20 uL cDNA to the wells ofa white plate (Roche 04729692001).
4. Add 1.6 uL of index **primer Fwd** 10 uM to corresponding cDNA according to the table

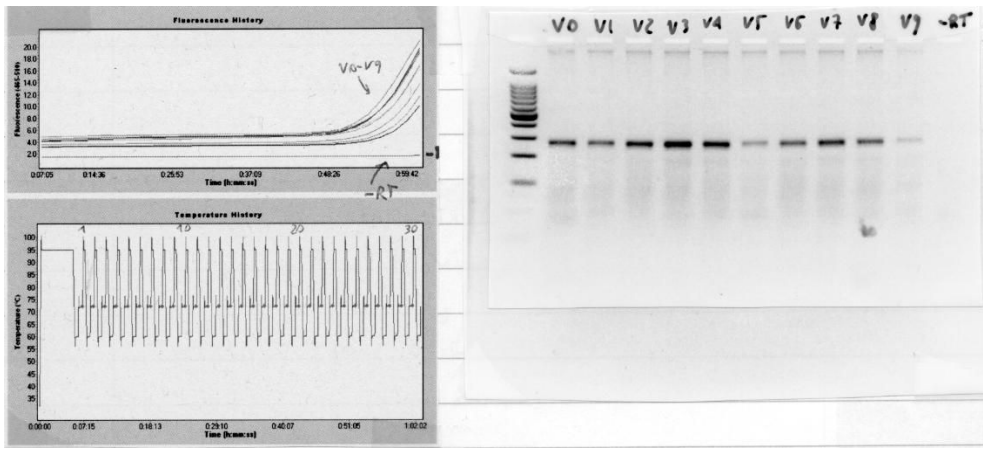
well	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
cDNA	V1	V2	V3	V4	V5	V6	V7	V8	V9	V0	no-RT
primer Fwd	37	39	40	41	74	75	76	78	79	36	36

5. Prepare the PCR mastermix

	Tot uL	Reactions	
	20	42	x

	1x	**Mix**	
2x SYBR Green (Roche)	10.5	441.0	uL
Primer **80 Rev**, 10 uM	0.42	17.7	uL
H2O	4.41	185.2	uL

6. Add 58.4 uL mastermix to each well containing 21.6 uL (cDNA + index primer) = 80 uL total.
7. Mix well and distribute 20 uL of the reaction in 4 different wells (rows A, B, C, D).
8. Seal and start qPCR machine. ~1hr later, abort the run when amplification is started (~cycle 30).
   Do not abort during denaturation step: wait for annealing/extension time to stop.
9. Repeat qPCR1 from step 3 with the remaining 1ug of cDNA template left.


Fluorescence History
V0-V9
-RT
Temperature History
V0 V1 V2 V3 V4 V5 V6 V7 V8 V9 -RT

Step 1.3 – **qPCR2**: Inserting the right Illumina end

Here we dilute the first PCR and use it as template for a second PCR. The forward primer is the same as the previous PCR. The reverse primer has a tail that introduces a docking site for the illumina.

Expected fragment 99 bp:

GACGCTCTTCCGATCTNNNNNNCAAGATCGCCGTGTAATAATbvhdhhhhbbvvddhhhhddTCTAGAGTCG
GGGCGGCATCGGAAGAGCGGTTCAGCA

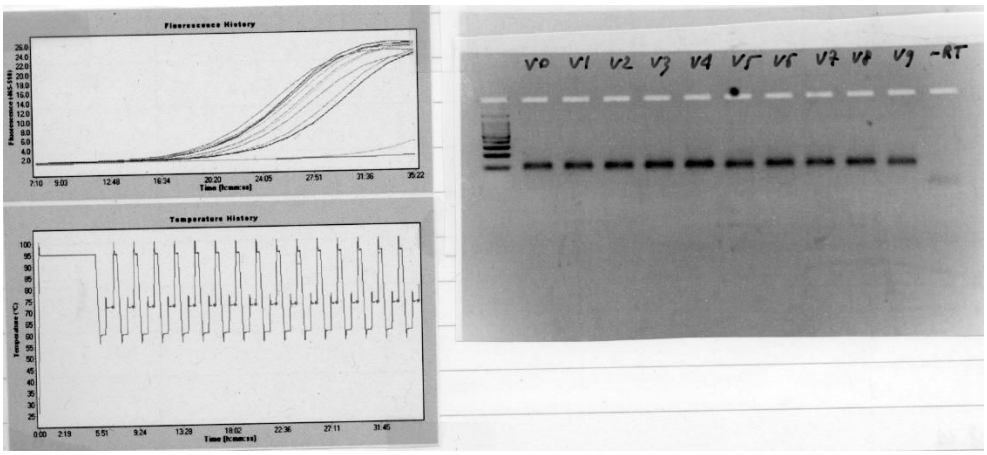
1. For each sample, pool the 8 qPCR1 together. Mix well, spin down.
2. In a second tube, dilute the qPCR 1:256 (3 uL PCR + 765 uL Tris 10 mM pH 8.0). Mix well.
3. The diluted PCR is used as template for the qPCR 2: distribute 10 uL diluted qPCR1 to the wells of a white plate (Roche 04729692001).
4. Add 0.8 uL of index **primer Fwd** 10 uM to corresponding cDNA according to the table

well	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
qPCR1	V1	V2	V3	V4	V5	V6	V7	V8	V9	V0	no-RT	H2O
primer Fwd	37	39	40	41	74	75	76	78	79	36	36	36

5. Prepare the PCR mastermix

		Tot uL	Reactions	
		20	24	x
		1x	**Mix**	
2x SYBR Green (Roche)		10.5	252.0	uL
Primer **42 Rev**, 10 uM		0.42	10.1	uL
H2O		4.41	105.8	uL

6. Add 29.2 uL mastermix to each well containing 10.8 uL (qPCR1 + index primer) = 40 uL total.
7. Mix well and distribute 20 uL of the reaction in 2 different wells (rows A, B).
8. Seal and start qPCR machine. Abort the run when amplification is started (~cycle 10).
   Do not abort during denaturation step: wait for annealing/extension time to stop.


Fluorescence History
Temperature History
V0 V1 V2 V3 V4 V5 V6 V7 V8 V9 -RT

Step 1.4 – **qPCR3**: Inserting universal Illumina tails.

This is the last step: all the previous qPCR2 are pooled, diluted and used as template for a third PCR that introduces Illumina tails. The PCR is gel purified and sent to sequencing.

Expected fragment: 180 bp

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNCAAGAT
CGCCGTGTAATAATHHDDDDHHBBVVDDDDHDBVTCTAGAGTCGGGGCGGCATCGG**AAGAGCGGTTCA**
GCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTTG

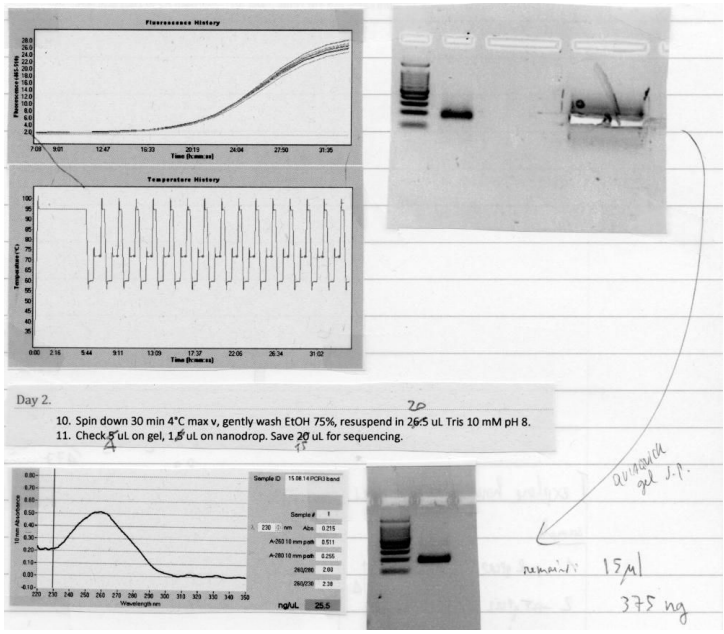
Day 1.

1. Pool 10 uL each of qPCR2 (samples V0 to V9). Do not pool –RT controls.
2. In a second tube, dilute the pool 1:256 (3 uL PCR pool + 765 uL Tris 10 mM pH 8.0).
3. The diluted PCR is used as template for the qPCR 3. We run 12 identical PCR reactions to prepare more material: distribute 5 uL diluted qPCR2 to 12 wells of a white plate (Roche 04729692001).
4. Prepare the PCR mastermix

		Tot uL	Reactions	
		20	12	x
		1x	**Mix**	
2x SYBR Green (Roche)		10.5	126.0	uL
Primer **44 Fwd**, 10 uM		0.42	5.04	uL
Primer **45 Rev**, 10 uM		0.42	5.04	
H2O		4.41	52.9	uL

5. Add 15 uL mastermix to each well containing 5 uL template DNA (20 uL total).
6. Seal and start qPCR machine. Abort the run when amplification is complete (~cycle 15). Do not abort during denaturation step: wait for annealing/extension time to stop.
7. Pool the PCR reactions together. Load 100 uL on a 2% gel.
8. Run on a gel, gel purify. Never expose to UV, use dark-reader.

9. If purifying with QUIAquick, elute in 50 uL $T_{10}E_{01}$, add 5.5 uL Na Acetate 3 M pH 5.2 and 140 uL EtOH, precipitate overnight -20°C.



Day 2.

10. Spin down 30 min 4°C max v, gently wash EtOH 75%, resuspend in 26.5 uL Tris 10 mM pH 8.
11. Check 5 uL on gel, 1.5 uL on nanodrop. Save 20 uL for sequencing.
12. Complete the fields of the genomic facility LIMS following to the template as example.

Lab/PI	Ueli Schibler
Submitter	Gianpaolo Rando
Name	Lib3_U2OS_Vb
Short Name	Lib3
Project	STARpromBC
Organism	Unknown
Starting Material Type	PCR clone
Starting Material Description	luciferase cDNA obtained after transfection of a library of barcoded luciferase plasmids
Library Protocol	Custom
Library Date	2014-08-15
Adapters	Custom
Multiplex Index	Custom
5prime Index	Custom
Fragment Size (Min)	180
Fragment Size (Max)	200
Concentration [ng/µl]	25
Quantification method	Nanodrop
Volume [µl]	15
Comment	custom library protocol using indexed PCR primers on the cDNA: see lib1 of 2014-01-24

Request(s)

Submitter	Gianpaolo Rando
Request Date	2014-08-21 10:25:22
Run type requested	multiplexed single read
Read length requested	100
Nb lanes requested	1
Multiplex level	1
Eland Alignment	No
Comment	custom library protocol using indexed PCR primers on the cDNA: see lib1 of 2014-01-24
Request done	No

Step 2 : Bioinformatics Analysis: Galaxy use

Day 1

1. When sequencing raw data is ready, do not wait to log into 'http://uhts-gva.vital-it.ch/lims' [Login: ------; Password: ----------] and go to 'My lab data' to readily copy the link pointing to the fastq.gz file.
2. Log into your usegalaxy.org account. Upload the fastq file link: go to **Get Data-Upload File** and paste the URL link, http://uhts-gva.vital-it.ch/symlink/Libn_nnnnnnnnnnn_nn_nn_nnn.fastq.gz copied from the UHTS-LIMS university site. Indicate FASTQ as file format. (1-2h)
3. Run on this fastq File a **Fastq Groomer** to convert options relating to the Fastq quality formats. The input Fastq quality scores type is Sanger&illumina 1.8+. Processing it will take an overnight. When started it is possible to turn off the computer (all the work is done on the Galaxy server).

Day 2

4. Convert this FastQ Groomer file (data2) to FastA with **FASTQ to FASTA converter**

5. Run **Trim sequences** on this FastQ to FastA (data3). First base to keep: 1; Last base to keep: 55.

6. Run **Clip** on the trim file. This tool clips adapters from the 3'-end of the sequences Starting and remove everything from the left.

Input Parameter	Value
Library to clip	4: Trim sequences on data 3
Minimum sequence length (after clipping, sequences shorter than this length will be discarded)	30
Source	user
Enter custom clipping sequence	GTAATAAT
enter non-zero value to keep the adapter sequence and x bases that follow it	20
Discard sequences with unknown (N) bases	No
Output options	Output only clipped sequences (i.e. sequences which contained the adapter)

7. Upload barcodes library file. **Get Data-Upload File.** Download barcodes_lib3.txt

1	2
v0	ACATCGCAAGATCGCC
v1	GCCTAACAAGATCGCC
v2	CACTGTCAAGATCGCC

v3 ATTGGCCAAGATCGC
C

v4 GATCTGCAAGATCGC
C

v5 ATCACGCAAGATCGC
C

8. Use **Barcode Splitter** on **Barcode to use**: Barcode File and **Library to split**: Clip File

Input Parameter	Value
Barcodes to use	9: Galaxy5-barcodes_lib4.txt
Library to split	6: Clip on data 4
Barcodes found at	Start of sequence (5' end)
Number of allowed mismatches	3
Number of allowed barcodes nucleotide deletions	3

File obtained:

Barcode	Count	Location
V6	8649154	Clip_on_data_4__V6.txt
V7	11152893	Clip_on_data_4__V7.txt
V8	14475321	Clip_on_data_4__V8.txt
V9	16006742	Clip_on_data_4__V9.txt
unmatched	1753682	Clip_on_data_4__unmatched.txt
v0	9021231	Clip_on_data_4__v0.txt
v1	2692041	Clip_on_data_4__v1.txt
v2	17467496	Clip_on_data_4__v2.txt
v3	15061998	Clip_on_data_4__v3.txt
v4	9852833	Clip_on_data_4__v4.txt
v5	14749844	Clip_on_data_4__v5.txt
total	120883235	

9. Upload each BC splitters results on your computer by clicking right on the link and select download link target.
10. Paste directly each link “clip on data” (click right on the link in the tab and choose “copy the link”) into **Get Data-Upload File** in the URL window, indicating Fasta as a format.

Day 3

11. Use **Reverse-Complement** for each of these files (10). You can use the multiple datasets tool
12. Run **Trim sequences** on each reverse complement file to cut the first 20 pb. You can use the multiple datasets tool

Input Parameter	Value
Library to clip	31: Reverse-Complement on data 21
First base to keep	1
Last base to keep	20

13. Run **Collapse** on each trim file
14. Download each Fasta File in a dedicated folder.
15. In MacOSX, open terminal in the dedicated folder and run the following commands

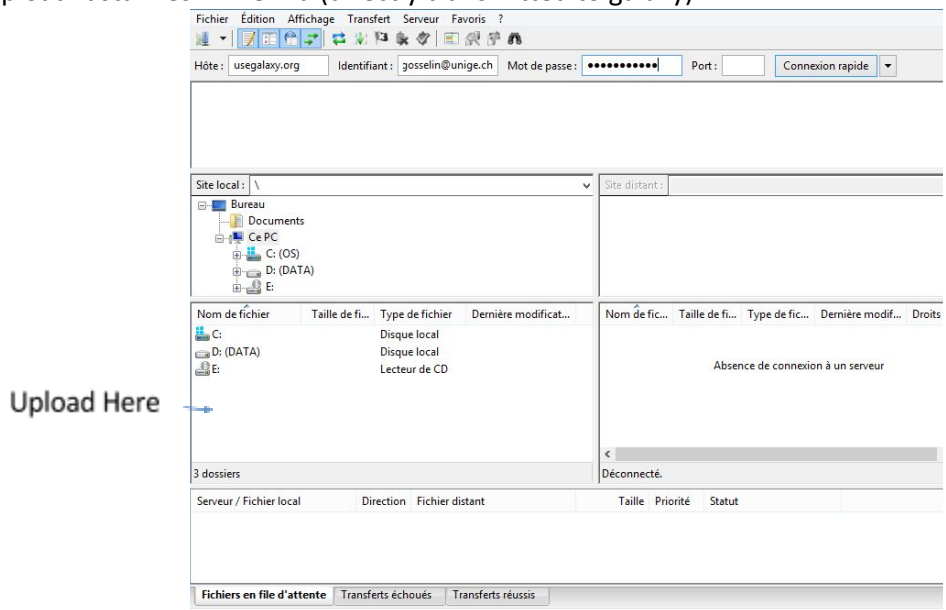
```
sed  -i "" 's/>/>v0-/g' v0.fasta
sed  -i "" 's/>/>v1-/g' v1.fasta
sed  -i "" 's/>/>v2-/g' v2.fasta
sed  -i "" 's/>/>v3-/g' v3.fasta
sed  -i "" 's/>/>v4-/g' v4.fasta
sed  -i "" 's/>/>v5-/g' v5.fasta
sed  -i "" 's/>/>v6-/g' v6.fasta
sed  -i "" 's/>/>v7-/g' v7.fasta
sed  -i "" 's/>/>v8-/g' v8.fasta
sed  -i "" 's/>/>v9-/g' v9.fasta

sed  -i "" 's/-/ /g' *
```

(the first set will replace > with >v0 to v0.fasta file, etc; the last row will replace - with a space) This step is required to convert fasta file to tabular.

16. Upload each Fasta file(FTP) with Filezilla. In FileZilla, fill name and password with galaxy login, host is usegalaxy.org.

Upload fasta files in filezilla (directly transmitted to galaxy)



17. Run **Fasta-to-Tabular** on each uploaded file. Use the run tool in parallel across multiple datasets. Enter all the file numbers in “convert these sequences”.

Input Parameter	Value
Convert these sequences	62: J0 Fasta.fasta
How many columns to divide title string into?	3
How many title characters to keep?	0

18. Run **cut-on-data** to remove column 2 (redundant). (cut columns: c1,c3,c4) for each file

19. Use **Join two datasets** to combine each file in the same tab. Always join the “join file” first with the cut on data file

Parameters:

Input Parameter	Value
Join	99: Join two Datasets on data 82 and data 98
using column	3
with	83: Cut on data 73
and column	3
Keep lines of first input that do not join with second input	Yes
Keep lines of first input that are incomplete	Yes
Fill empty columns	no_fill

Join cut on data vo+ cut on data v1 → file A
Join file A+ cut on data V2 → File B
Join File B+ cut on data V3 → File C
.....
Join File I+ cut on data V9 → File J

20. Run **cut-on-data** to remove all the redundant columns (BC) and the index columns. Select columns in the right order to have in the tab reads going from IN0 to IN9
    IE:

Input Parameter	Value
Cut columns	c3,c5,c2,c8,c11,c14,c17,c20,c23,c26,c29
Delimited by	Tab
From	105: Join two Datasets on data 88 and data 104

21. Run **select first** to select the number of line you want to keep.

Step 3 : Bioinformatics Analysis: STATISTICS