umcg
Genetica
Genome Analysis Facility

Title:	Guidelines Analysis of RNA Quantity and Quality for Array-Based Illumina Expression Studies
Protocol number:	GAF E001
Version:	Version 3
Date:	June 15th, 2014
Author:	P. van der Vlies, C. C. van Diemen

Guidelines Analysis of RNA Quantity and Quality for Illumina Array-based Expression Studies	GAF E001

Introduction

To ensure good quality data and a successful and rapid processing of total RNA samples for Illumina Array-based Expression Studies, there are a number of points to consider. Also the RNA quality has to fulfil a number of requirements to be suitable for use. Properties such as purity, integrity and concentration are of great importance, because they can have a significant impact on the experiment's results.
All the required information before start of an Illumina Array-based Expression Study, performed by the Genome Analysis Facility of the UMCG (GAF), is described in this document.

Applications and limitations of the technique

The Illumina Expression BeadChips can provide genome-wide transcriptional coverage of well-characterized genes, gene candidates, and splice variants for human and mouse tissues. These data can be used to answers multiple research questions such as::

- Identification of **genome-wide differential expression of genes** across two or more conditions with statistical significance and/or biological significance (e.g. fold change)
- **Pathway analysis** (a.k.a. gene set analysis) to infer correlation of differential expression evidence in the data with pathway knowledge from established databases
- **Co-expression analysis** and network analysis of RNAseq data are used to investigate potential transcriptional co-regulation and gene interactions
- **Prediction analysis** for clinical practice

Points to consider prior to the start of the experiment

Practical aspects

- RNA isolation and quantification: Use the same method throughout study to prevent clustering of samples per method.
- Avoid clustering of samples due to batch differences or technical effects. Randomize the samples, but prevent sample mix-ups.
- Biological replicates are necessary, technical replicates not.

Limitations

- Only poly-A tailed RNA transcripts are assessed.
- Polymorphisms affect hybridization intensity and may introduce bias.
- Contamination with e.g. abundance of globin mRNAs in whole blood, significantly influence the whole gene expression profile.
- Resolution of arrays is limited for very low or high expressed genes.
- No identification of novel splicevariants.

design of the study and data analyses

- Resource limitation vs. statistical power: ensure sufficient samples to allow robust results
- The technique is meant for genome-wide expression analysis, not to analyse only a few genes. qPCR is then better suited.
- Replication problems between studies may occur due to diversity of study designs
- Cellular diversity within samples: expression variation is largely within or among cells
- Normalization of the data is necessary
- No consensus on statistical thresholds for differential expression
  - fold-change
  - Bonferroni correction
  - false discovery rate

RNA isolation

RNA isolated by most isolation methods usually meets the criteria for a successful next-generation sequencing experiments. The RNA should be dissolved in 15 µl Elution buffer or 10 mM Tris-HCL pH 8.0, 1 mM EDTA. We have successfully processed RNA obtained by the Qiagen RNeasy-kit, Norgen kits or Trizol reagent.

- The GAF guarantees a successful experiment for samples containing at least 300 ng pure and non-degraded RNA.
- Treat all samples in the same way as much as possible to avoid batch-effects.

RNA quantity and quality

1. The quantity and purity of the isolated DNA should be initially verified with a spectrophotometer, ideally by the Nanodrop spectrophotometer.
   The GAF quantity and purity criteria for RNAseq are:
   a. Concentration should be > 20 ng/µl. (measurement with capillary electrophoresis)
   b. Total volume > 15 µl
   c. The OD 260/280 ratio should be above 2.0, the OD 260/230 ratio should be above 2.0 (Nanodrop measurement)
   d. RQI/RIN-score is > 7.0. (measurement with capillary electrophoresis)
2. The measurement of the integrity of the RNA and a more precise measurement of the concentration should be performed using capillary electrophoresis (Bio-Rad Experion Bioanalyzer, Agilent Bio-Analyzer or PerkinElmer LabChip GX)device.
3. As extra service, the GAF can perform the nanodrop measurement and capillary electrophoresis for RNA samples. This is done in a weekly routine according to document GAF ALG005

Preparing your samples and shipping them

1. At least 50 µl of the RNA solution should be transferred to 96-well plates (for instance: Greiner Bio-One, PCR-plate full skirted, order nr. 652270).
2. To avoid seeing batch differences in the final data, it is important to put the samples in the plate in a random order, starting with position A1, then B1, etc. Do not leave any blank wells between samples.

3. The 96-well plates should be firmly closed, preferably with a Thermo-seal (*Thermo scientific order no. AB-0559*) using a heat-plate sealer. Alternatively, if a heat-plate sealer is not available, adhesive PCR foil seals (*Thermo scientific order no. AB-0626*) can be used.
4. Record the sample IDs in a special XLS file according to the instructions below. Use The facility doesn't accept samples with names of individuals. A unique plate name should be given to each 96-well plate, using a permanent marker to annotate each plate on both sides.
5. The plates should be stored and transported at –80°C. Deliver the samples on dry ice to the Genome Analysis Facility lab or send your samples by courier to the Genome Analysis Facility laboratory at the Department Genetics, UMCG. The shipment should be placed in boxes containing a sufficient amount of dry ice.
6. To avoid any delay during the weekend, please ship the material at the beginning of the week.
7. The Genome Analysis Facility is not responsible for long time storage of samples or intermediate products. Left overs are thrown away after finishing the project.

Documenting the sample and plate layout

The administration of all sample IDs and their corresponding wells in the 96-well plate should be recorded in an XSL file (Excel) (see figure 1). This XSL file should summarize all the required information:

- Number
- Plate name
- Well with corresponding unique sample ID
- Information on the RNA quality: RNA concentration, OD 260/280 ratio, OD 260/230 ratio.

nr	Plate	Well	SampleID	RNA-concentration (ng/ul)	OD260/280	OD 260/230	Volume (ul)	RIN/RQI value
1	UMCG-1	A1	345	71	2.12	2.39	20	8
2	UMCG-1	B1	456	82	2.07	2.4	20	8
3	UMCG-1	C1	564	56	2.12	2.38	20	9
4	UMCG-1	D1	873	65	2.09	2.28	30	10
5	UMCG-1	E1	937	87	2.1	2.34	30	8
6	UMCG-1	F1	345	93	2.06	2.23	20	9
7	UMCG-1	G1	244	52	2.11	2.43	20	10
8	UMCG-1	H1	879	79	2.12	2.19	25	8
9	UMCG-1	A1	435	80				
	UMCG-1	B2						

Figure 1. Example of the attached XSL file

Please email the XLS file with your sample information to: p.van.der.vlies@umcg.nl

Delivery address

Genome Analysis Facility
Eriba building, 5th floor
Dept. Genetics UMCG CB53
Hanzeplein 1
9713 GZ Groningen
The Netherlands

Phone number: +31 (0)50-3617100
Before submitting your samples, please contact the Genome Analysis Facility