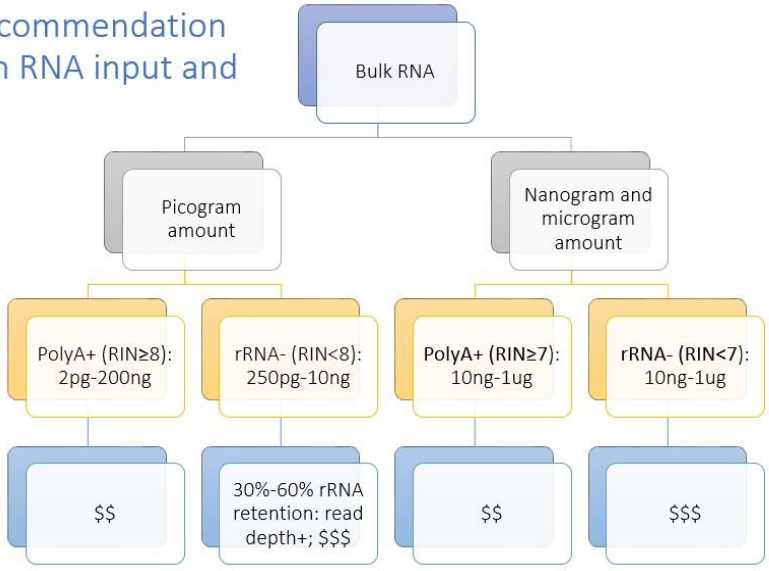
Princess Margaret Genomics Centre
101 College Street
PMCRT, Rm 9-601A
Toronto, ON, M5G 1L7

PMGC SAMPLE SUBMISSIONS GUIDELINES FOR BULK RNA

This document covers FAQ about RNA-Seq as follows:

Assay recommendation based on RNA input and quality

This approach is only appropriate for gene expression profiling experiments that are looking for a quick snapshot of highly expressed genes.



Nanogram and microgram amounts are appropriate for experiments looking for a quick snapshot of highly expressed genes, a more global view of gene expression, information on alternative splicing and for experiments looking to get an in-depth view of the transcriptome (incl. low-expression genes), or to assemble new transcripts.

*Minimum input amounts are required for optimal kit performance but may not be appropriate for the biological question you are asking. Depending on RNA quality and/or experiment requirements, the *recommended input amounts* can vary greatly from *minimum input amounts* required by the kit.*

Read Depth

Highly expressed genes	Global view of Gene expression + alternative splicing	In-depth view of transcriptome + transcriptome assembly
• 5-25M clusters	• 30-60M clusters	• 100-200M clusters

Julissa Tsao
For Bulk RNA inquiries,
(416) 581-7439
Julissa.Tsao@uhn.ca

Dr. Troy Ketela, Head of Operations
For new project inquires,
(416) 634-8816
Geneservice@pmgenomics.ca

Princess Margaret Genomics Centre
101 College Street
PMCRT, Rm 9-601A
Toronto, ON, M5G 1L7

1. RNA Input Amounts and RNA Quality

RNA Input Amount

- The minimum RNA input amount required for your project will depend on your experimental design, data requirements, and RNA quality.
- The minimum input amount required for optimal kit performance may not be appropriate for the biological question you are asking. Depending on RNA quality and/or experimental requirements, the ***recommended input amounts*** can vary greatly from ***minimum input amounts*** required by the kit.
- Broadly, the more involved your analysis the higher the recommended RNA input amount. Higher RNA inputs ensure increased library complexity and improved transcript coverage.
    - If you have **picogram amounts** of RNA:
        - *This approach is appropriate for gene expression profiling experiments that are looking for a quick snapshot of highly expressed genes.*
        - We offer a polyA+ based picogram input amount library prep chemistry for high quality intact total RNA. RNA Integrity Number (RIN) must be ≥8.
        - We also offer an rRNA depletion-based picogram protocol for low quality (partially degraded, highly degraded, and FFPE) RNAs, however rRNA retention/contamination ranges from 30%-60% with this chemistry and has to be dealt with bioinformatically. This approach requires deeper sequencing and a larger budget.
    - If you have **nanogram or microgram amounts** of RNA:
        - *Nanogram and microgram amounts are appropriate for experiments looking at global gene expression, information on alternative splicing, in-depth view of the transcriptome (incl. low-expression genes), to assemble new transcripts, and for a snapshot of highly expressed genes.*
        - Most chemistries that deal in this input range, call for a minimum of 10 ng total input amount. *The minimum amount required for optimal kit performance but may not be appropriate for the biological question you are asking.*

RNA Quality

- We strongly recommend that you **use RNase inhibitors** throughout your experiment wherever possible as well as in your final elution buffer (spike in RNase Inhibitors into the elution buffer prior to elution). This is especially important if you are working with tissues known for high RNase content.
- **RNA samples must be free of genomic DNA** (gDNA). gDNA can skew RNA quantification and normalization leading to noisy data and batch effects. Please see *4. Sample Submission Requirements* for more info.
- FFPE samples have a **significant variability in library performance**. FFPE samples contain chemical modifications that decrease the efficiency of reverse transcription, leading to reduced overall cDNA yield and ligation inefficiencies. The age of FFPE samples and the extent of RNA damage plays a direct

Julissa Tsao
For Bulk RNA inquiries,
(416) 581-7439
Julissa.Tsao@uhn.ca

Dr. Troy Ketela, Head of Operations
For new project inquires,
(416) 634-8816
Geneservice@pmgenomics.ca

Princess Margaret Genomics Centre
101 College Street
PMCRT, Rm 9-601A
Toronto, ON, M5G 1L7

role in how accessible RNA is to NGS enzymes. This complicates the normalization of input amounts, as it is challenging to predict the proportion of accessible RNA. Lower input amounts exacerbate data variability and lead to lower quality libraries, lower mapping rates, higher duplication rates, and lower proportion of usable reads.

2. Assay Recommendation

Prior to making a recommendation we take into consideration four factors:

1. What are you looking for in your data? For example,
    a. Are you interested in coding or also non-coding transcripts?
    b. Are you looking for low-expression genes and/or rare events, or are you assembling a new transcriptome?
    c. Are you interested in directional (strand-specific) or non-directional data?
2. The quality, purity and amount of RNA available.
3. The number of samples.
4. Your budget.

Based on your project requirements we recommend one of the following approaches:

1. Full-length directional or non-directional coding RNA-Seq **(mRNA-Seq / polyA+ enrichment)**
- This method depends on an intact 3' polyA tail and thus necessarily requires high quality intact total RNA as input, usually with a RIN≥7, depending on library prep chemistry used.
- Of the two assays, this is a more cost-effective approach, both in terms of library preparation and sequencing read depth, however it necessitates a RIN ≥8.

2. Full-length directional or non-directional coding and non-coding RNA-Seq **(total RNA-Seq / ribo-depletion)**
- This method depends on rRNA depletion probes and is recommended for samples with a RIN of <7. That includes partially and highly degraded RNA samples and FFPEs.

3. Read Depth Recommendation

Read depth varies depending on the goals of the RNA-Seq study. Most experiments require 5–200 million clusters per sample, depending on organism complexity and size, along with project aims.

- Gene expression profiling experiments that are looking for a quick snapshot of highly expressed genes may only need 5–25 million clusters per sample.
- Experiments looking for a more global view of gene expression, and some information on alternative splicing, typically require 30–60 million clusters per sample.

Julissa Tsao
For Bulk RNA inquiries,
(416) 581-7439
Julissa.Tsao@uhn.ca

Dr. Troy Ketela, Head of Operations
For new project inquires,
(416) 634-8816
Geneservice@pmgenomics.ca

- Experiments looking to get an in-depth view of the transcriptome, or to assemble new transcripts, may require 100–200 million clusters.
- Targeted RNA expression requires fewer reads. For example, Illumina recommends 3 million clusters per sample for TruSight RNA Pan Cancer and TruSight RNA Fusion Panel.
- miRNA-Seq or small RNA Analysis experiments may require even fewer reads than whole transcriptome sequencing. This requirement varies significantly depending on the tissue type being sequenced. Illumina strongly recommends using the primary literature to determine how many reads are needed, with most applications ranging from 1–5 million reads per sample.

Source: Considerations for RNA-Seq read length and coverage (illumina.com)

4. Sample Submission Requirements

- Please complete the **PMGC Sample Submission Form**. This form is uploaded to our LIMS and it is used to track your project through our pipeline. If you have sent/submitted your samples to our facility without this form, your samples will be stored appropriately and the project will be placed on hold.
- Please submit any QC data (TapeStation, Bioanalyzer, NanoDrop, Qubit) you may have collected for your samples prior to shipping/submitting your samples to our core facility.
- Please note that we do not guarantee library performance for low-quality, low-input and 'dirty' samples:
    - RNA samples must be of pure quality, i.e. have clean 260/280 (~2.0) and 260/230 (~2.2) ratios.
    - Samples must be free of salts (e.g., Mg2+, or guanidinium salts, divalent cation chelating agents (e.g., EDTA or EGTA) or organics (e.g., phenol or ethanol). Please consider running an additional clean up column if your samples are especially dirty.
    - Samples must be free of genomic DNA (gDNA).
        - In RNA-Seq applications, RNA must be free of DNA. Random hexamers and short oligo(dT) primers used in RNA-seq applications cannot distinguish between RNA and DNA molecules and will hybridize to residual genomic DNA. Additionally, reverse transcriptases are able to use DNA as the template molecule. As a result, genomic DNA will carry through the entire RNA-Seq workflow. This can cause biases and quantification issues during the final data analysis steps. Therefore, it is critical to remove any residual genomic DNA to obtain the best quality data.
            - There are three options available for genomic DNA removal:
                1. On-column DNase digestion (during extraction)
                2. Acid Phenol Extraction of RNA (extraction method)
                3. DNase I digestion (post extraction)
        - We do not offer DNase treatment services at this time. RNA samples must be submitted to us without genomic DNA. Our internal intake QC pipeline can help identify residual

gDNA in intact RNA samples, however we have no way of identifying DNA contamination in degraded and FFPE RNA samples. DNA contamination is highly likely to be an issue in cells that have a higher DNA:RNA ratio (cell lines, blood cells, bacteria) and FFPE samples due to cross-linking of macromolecules.

- As a general rule on elution volumes, the higher the expected RNA yield post extraction, the more volume is acceptable. For highly concentrated samples, please keep the volume around 50uL. For samples with low RNA yields, please keep the volume around 20uL. Please elute in RNase-free water. If you are especially worried about degradation, you can spike RNase Inhibitor(s) into the RNase-free water prior to elution. Please indicate on the PMGC Sample Submission Form that elution water contains RNase inhibitor(s).

Sample Drop-off / Shipping

If dropping off samples: Please **schedule your drop off date and time in advance** with your PMGC contact person.

- Your PMGC contact will meet you at the **9th floor elevator lobby** of the Princess Margaret Cancer Research Tower (PMCRT) at your pre-arranged time. PMCRT is the East Tower of the MaRS building, near the corner of College and Elizabeth Street entrance.
- Email or call/text when you are at the designated meeting area and your PMGC contact will come to collect the samples.
- REMINDER: Transport samples using appropriate means of storage (e.g. on dry ice for frozen samples). Please confirm with PMGC if any questions.

If shipping samples: Please ship out on **Monday/Tuesday** to prevent weekend delays. Place a generous supply of dry ice to ensure dry ice will remain for the duration of the delivery time. For international clients, we recommend shipping with World Courier. Within Canada, or if shipping DNA/RNA, we recommend FedEX Next Day Priority services.

Shipping address:

Attn: (insert PMGC contact person)
Princess Margaret Genomics Centre
101 College St.
PMCRT, Rm 9-601A
Toronto, Ontario M5G 1L7
Canada

Julissa Tsao
For Bulk RNA inquiries,
(416) 581-7439
Julissa.Tsao@uhn.ca

Dr. Troy Ketela, Head of Operations
For new project inquires,
(416) 634-8816
Geneservice@pmgenomics.ca