Development of novel pipeline for miRNAs, benchmarking and comparing with clinical CLL NGS data

Advisors:
Dr. Sriram K.
Dr. Anubha Gupta
Dr. Ritu Gupta

Presented by:
Vivek Ruhela
Ph.D. 18202 (CB)

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Outline

- Background
- Motivation
- Problem Statements
- Role of small non-coding RNAs in CLL and their impact on clinical outcome
  - Workflow for the identification of clinically relevant biomarkers in CLL
  - Dysregulated miRNAs in CLL
  - Impact of dysregulated miRNAs on the clinical outcomes in CLL

Background

- Blood cancer is a hematological malignancy caused by neoplastic proliferation of malignant blood cells.
- Blood cancer can be further categorized into the following groups:
  - Leukemia
  - Multiple Myeloma (MM)
  - Lymphomas
- Leukemia is a cancer in which malignant cells are found within bone marrow (BM) and blood cells (BC).
- Leukemia can be further classified into two groups based on the rate of onset of disease as follows:
  - Acute or chronic
  - Myeloid or lymphoid (also known as lymphocytes)
- Multiple myeloma is a malignancy in which there is a abnormal proliferation of malignant plasma cells (PC).
- Lymphomas is a malignancy in lymphoid lineage and causing a lymphomatous static tumor.

Classification of snc/mnc-RNA's

S. No	Name	Seq length (nt)	Biological Functions	Example
1	miRNAs	17 nt-24 nt	Modulation of mRNA Expression	hsa-miR-155, hsa-miR-423 etc.
2	Isomirs	22 nt (avg) [difference in ref mature seq]	Functionally cooperative partners of canonical miRNAs	hsa-let-7e-5p, hsa-let-7e-3p etc
3	Pi-RNAs	26-31	Silencing of transposable elements in germline cells; could function in somatic cells	hsa-piR-4987, hsa-piR-651
4	moRNAs	15 nt – 30 nt	miRNA co-product and alters gene expression and inhibits the proliferation of vascular smooth muscle cells	hsa-moR-21

S. No.	Name	Seq length (nt)	Biological Functions	Example
1	snoRNAs	41 nt-180 nt	Pre-rRNA processing; 2-*O*-methylation and pseudouridination of rRNAs	SNORD73, SNORA69 etc.
2	tRNAs and derivates	>= 72 nt	Play role as adaptors during translation of the genetic code. Gene expression reguation. the function of tRNA derivatives remains to be Explored	tRF-1, tRF-3, tRF-5, tRF-1001 etc.

Table-1: Classification of sncRNAs/mncRNAs

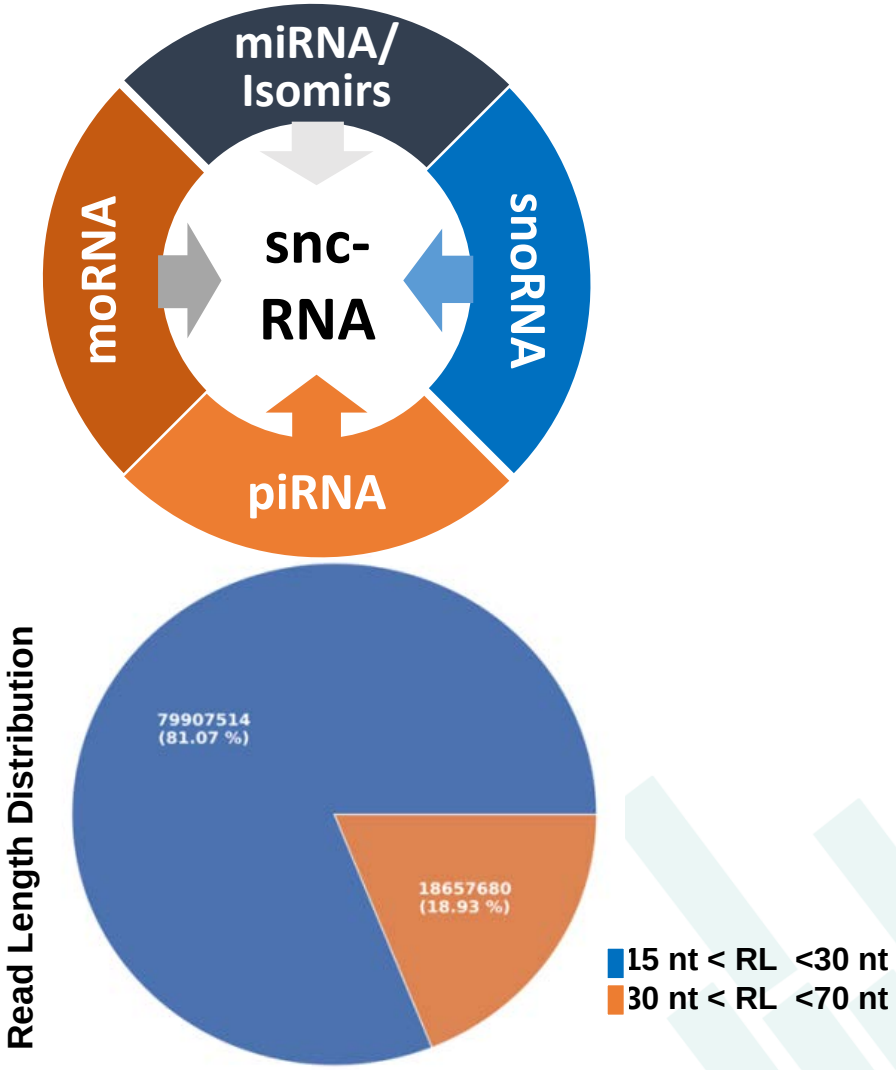


Figure-1: (A) Categories of sncRNA, (B) read length distribution

Challenges and Motivations

- According to National Cancer Institute, a cancer can be characterized by a biological molecule that is a sign of abnormal process or a condition of disease called as biomarkers [2].

- Identification of cancer biomarker is a challenging task due to several challenges such as
  - Extracting reliable information from huge amount of genomics data.
  - Vast computational resources are required for data processing and analysis.
  - Rigorous analytic and clinical validation must be established before a biomarker is used in the clinic.

- With increasing complexity of genomics data, artificial intelligence can be helpful to identify meaningful patterns and infer salient information from multi-omics datasets that can be utilized to halt the disease progression.

Problem Statements

1. What are the contribution of small non-coding RNAs (miRNAs, piRNAs, and snoRNAs) in disease pathogenesis of CLL?

2. Are the dysregulation of sncRNAs associated with any clinical prognostic factor and affect the patient survival?

Proposed workflow for sncRNA identification from RNA-Seq data

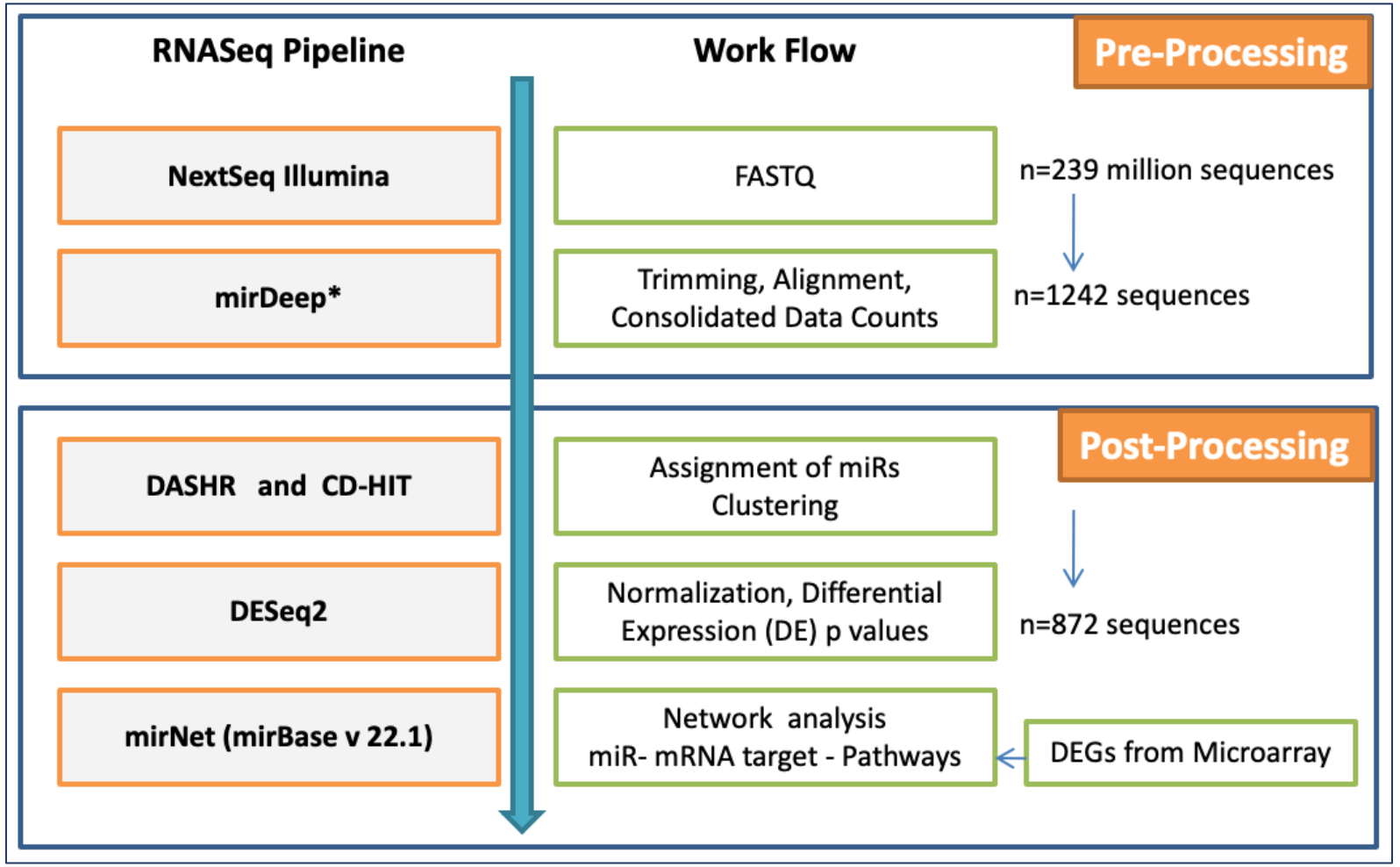


Figure-1: Workflow for RNA-Seq data analysis for biomarker identification in CLL

Dysregulated sncRNAs in CLL and Annotation of novel miRNAs

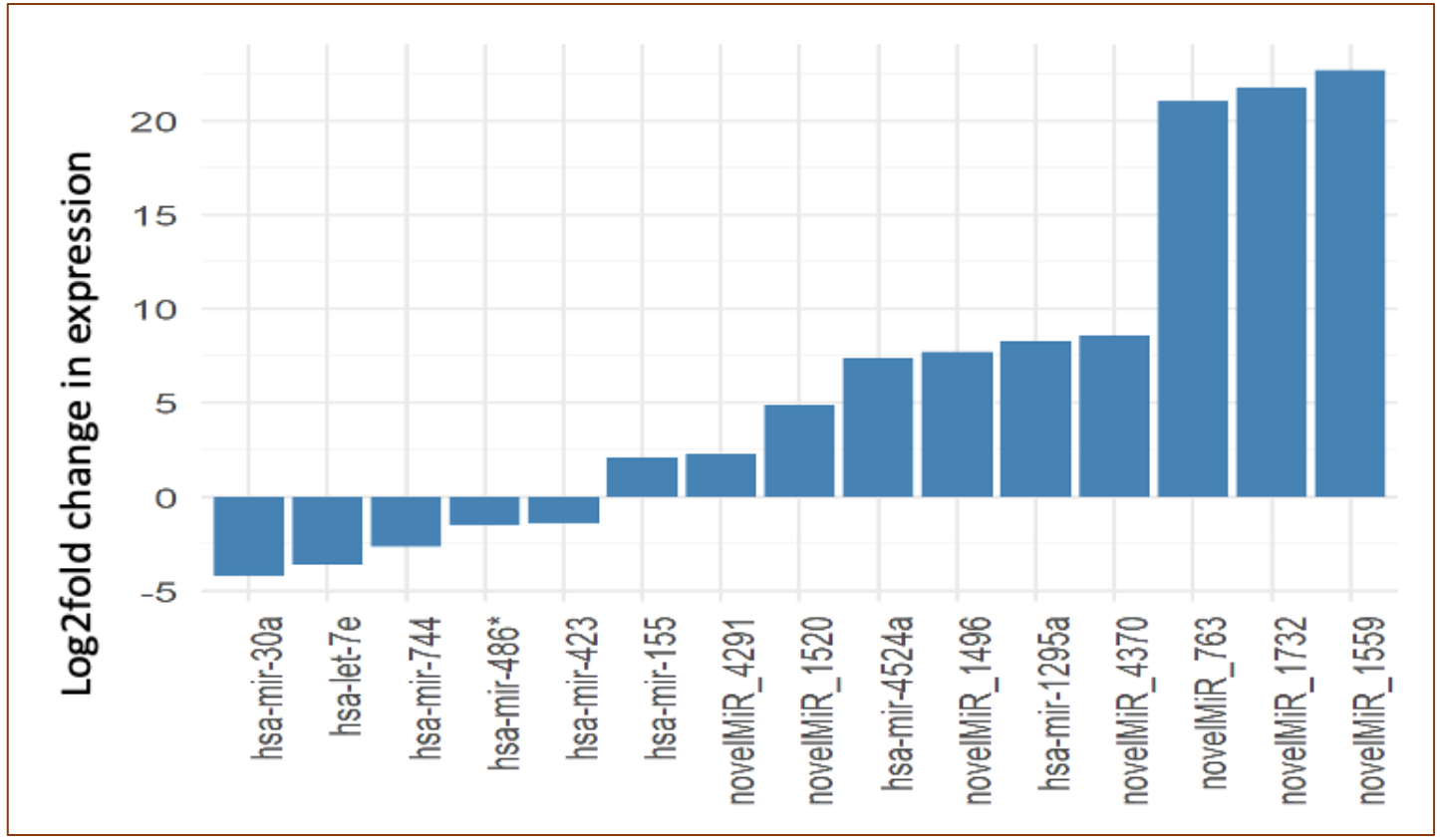
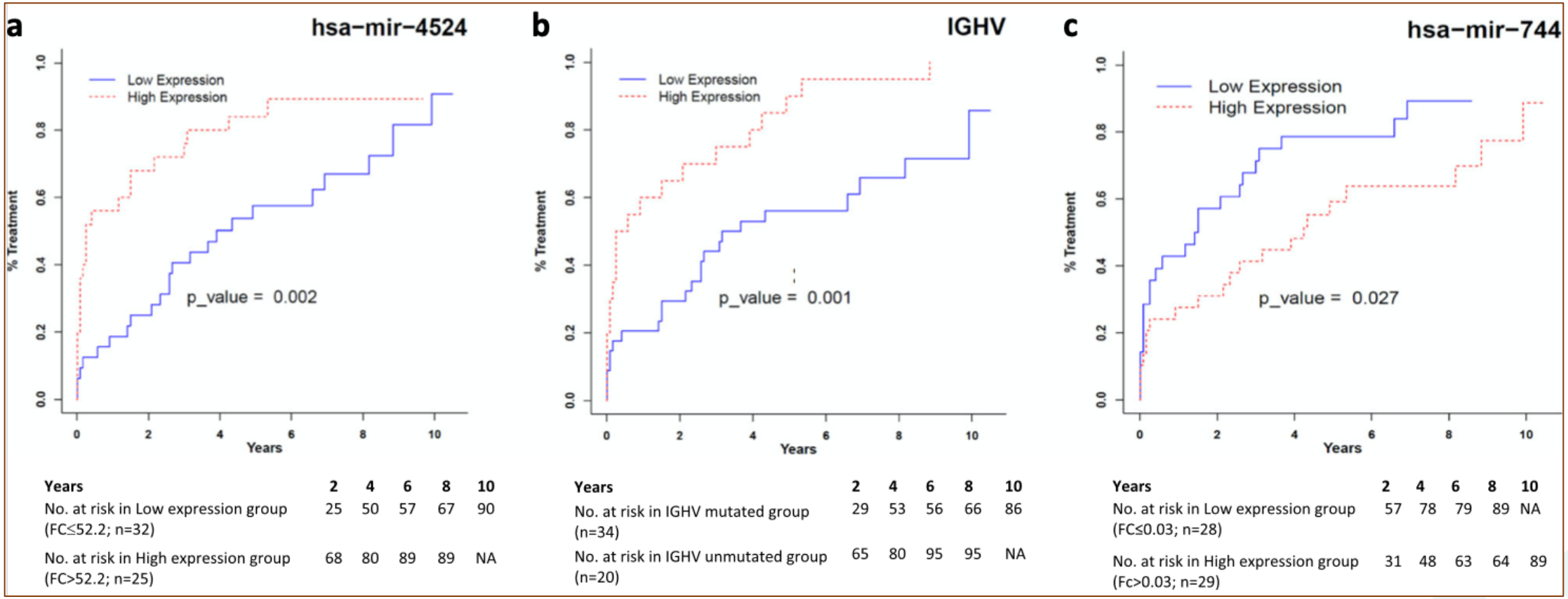


Figure-2: Dysregulated known and novel miRNAs in CLL (GSE123436)

Table-1: Annotation of dysregulated novel miRNAs using DASHR

Novel miRNAs	sncRNA identified based on homological similarity using DASHR database
novelMiR_4370	piR-36225
novelMiR_763	piR-30799 and snoRNA-U43
novelMiR_4370, novelMiR_1559, novelMiR_1732, novelMiR_4291, novelMiR_1520	tRNA molecules located on chromosomes 1, 6, 7, 9, 11, 12, 14, 15, 16, and 17.

Impact of miRNA dysregulation on clinical outcomes in CLL



Years	2	4	6	8	10
No. at risk in Low expression group (FC≤52.2; n=32)	25	50	57	67	90
No. at risk in High expression group (FC>52.2; n=25)	68	80	89	89	NA

Years	2	4	6	8	10
No. at risk in IGHV mutated group (n=34)	29	53	56	66	86
No. at risk in IGHV unmutated group (n=20)	65	80	95	95	NA

Years	2	4	6	8	10
No. at risk in Low expression group (FC≤0.03; n=28)	57	78	79	89	NA
No. at risk in High expression group (Fc>0.03; n=29)	31	48	63	64	89

Figure-3: Cumulative incidence plots demonstrating risk of treatment in CLL patients stratified on the basis of level of expression of a miR-4524a, b IGHV mutation status, and c miR-744. The cut-offs for defining low and high expression of miRNA and the number of cases in each subgroup are shown below the curves. p-values and hazard ratios as obtained in the Fine- Gray model of multivariate analysis is shown inside the curve.

Publication

Gurvinder Kaur, **Vivek Ruhela**, Lata Rani, Anubha Gupta, Krishnamachari Sriram, Ajay Gogia, Atul Sharma, Lalit Kumar, and Ritu Gupta. "RNA-Seq profiling of deregulated miRs in CLL and their impact on clinical outcome." Blood cancer journal 10, no. 1 (2020): 1-9.