

**Exploring the potential of small RNA cargo of extracellular vesicles as a next-generation
biomarker for B lymphoblastic leukemia**

By © Modeline Nicholas Longjohn, B.Sc. (Hons.), MSc (Res.)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Faculty of Science, Department of Biochemistry

Memorial University of Newfoundland

August 2023

St. John's

Newfoundland and Labrador

Abstract

Extracellular vesicles (EV) are novel promising sources of biomarkers in human diseases. Despite the promise that EVs hold, there are currently lots of unknowns, which limit our understanding. Therefore, the work presented in this thesis explored questions about EVs in pediatric B cell acute lymphoblastic leukemia (B-ALL) and compared the effect of collection tube treatment on EVs from selected human biofluids. For B-ALL, we wanted to identify a miRNA signature of pediatric B-ALL, which could be explored in EVs. We performed a reanalysis of miRNA raw data and metaanalysis of published miRNAs data; to identify differentially expressed miRNAs (DEmiRs). Neither the reanalysis nor the metaanalyses revealed a consensus miRNA signature. However, we identified promising miRNAs which could be further investigated. Next, we compared EVs from blood plasma of pediatric B-ALL patients and non-cancer donors (NCD) for size, concentration, and small RNA profile differences. We found that B-ALL plasma contains more EVs than NCD plasma. Furthermore, RNA-seq analysis of small RNAs revealed a signature of differentially packaged and exclusively packaged RNAs (including miRNAs, lncRNAs, mRNAs, and tRNAs) that distinguish NCD from B-ALL. Next, we investigated the small RNA cargo packaging patterns in EVs from pediatric B-ALL primary lymphoblasts and immortalized cell lines. We found that RNA biotypes packaged from cells into EVs include mature and precursor miRNAs and fragments of mRNA, lncRNA and tRNA. We also found that pediatric B-ALL primary lymphoblasts and immortalized cell lines package multiple RNA biotypes, using putative motifs shared by the selectively packaged RNAs. Finally, we compared EV characteristics (size, shape, concentration) and selected EV miRNA levels between non-Streck (EDTA for plasma) and Streck treated collection tubes for urine and blood sample collection. We found that neither treatment affected EV size and concentration in plasma or urine. However, Streck treatment led to reduced RNA

cargo in EVs, compared to non-Streck. Taken together, this work provides new insights into EVs in pediatric B-ALL, and blood and urine sample collection for EV profiling. This is knowledge that can facilitate next generation research for EV based biomarkers and monitoring disease burden.

General Summary

Extracellular vesicles (EV) are promising tools that could be used for diagnosing and monitoring diseases. Since their discovery, scientists continue to uncover details about how useful EVs can be. Despite the potential of EVs, there are still many unknowns about EVs. Therefore, the research presented in this thesis looks at some of these unknowns. Overall, my research looked at EVs in pediatric B cell acute lymphoblastic leukemia (B-ALL) and bodily fluids (urine and blood). For B-ALL, we looked for differences in miRNA content of cells from B-ALL patients and healthy people. Using computer-based analysis, we found that current methods could not sufficiently classify those differences. Next, we took blood from children that were healthy or had B-ALL, and compared the size, number of particles, and type of small RNAs in the EVs from both groups of samples. We found that the blood from children with B-ALL contained more EVs than healthy children. Also, we observed that the two sets of EVs had a group of small RNAs that distinguish disease from healthy. Next, we took cells from patient blood and investigated the pattern around how small RNA was packaged into EVs. We found that different types of small RNA were carried from cells into EVs. Finally, we collected blood and urine from donors (most of whom had cancer) in two types of collection tubes (each tube was treated with a different chemical). We isolated and compared EVs from both collection tubes. We found that the properties of EVs from both collection tubes were similar. However, in one collection tube, we had a reduced amount of RNA in the EVs. This suggests that both collection tubes did not alter the physical properties of the EVs, though EVs in one had a lower RNA content. Taken together, this work answers questions about EVs in pediatric B-ALL and how collection tubes affect EV properties. This is new knowledge that brings scientists one step closer to being able to use EVs for diagnosing and monitoring disease.

Table of Contents

Abstract	<i>i</i>
General Summary	<i>iii</i>
Acknowledgement	<i>viii</i>
List of Tables	<i>x</i>
List of Figures	<i>xi</i>
List of Abbreviation and Symbols	<i>xv</i>
List of Appendices	<i>xviii</i>
List of Supplemental Files and Figures	<i>xviii</i>
Chapter 1: Introduction	<i>1</i>
1.1 B Lymphocytes	<i>1</i>
1.1.1 Discovery of B cells.....	<i>1</i>
1.1.2 Hematopoiesis.....	<i>4</i>
1.1.3 Normal B cell development.....	<i>5</i>
1.1.4 Types of B cells	<i>13</i>
1.2 Gene expression for B cell development	<i>18</i>
1.3 Aberrant B cell development and hematological malignancies (HM)	<i>25</i>
1.3.1 Leukemia	<i>25</i>
1.3.2 B lymphoblastic leukemia	<i>26</i>
1.3.3 Pediatric B-ALL treatment	<i>34</i>
1.3.4 Pediatric B-ALL relapse and Measurable residual disease (MRD)	<i>36</i>
1.3.5 Other Hematological malignancies	<i>37</i>
1.4 Extracellular Vesicles (EVs)	<i>39</i>
1.4.1 EV secretion and cargo.....	<i>46</i>
1.4.2 EVs and Liquid biopsy.....	<i>51</i>
1.4.3 Knowns and unknowns about extracellular vesicles	<i>55</i>
1.4.4 EVs in the regulation of normal immune function	<i>55</i>
1.4.5 EVs in immune evasion	<i>57</i>
1.4.6 EVs and the bone marrow microenvironment.....	<i>59</i>
1.4.7 EVs in angiogenesis, motility, and adhesion	<i>63</i>
1.4.8 EVs in resistance to cancer therapies	<i>66</i>
1.4.9 Clinical applications for EVs in HM	<i>67</i>
1.4.10 Future Considerations	<i>71</i>
1.5 Research hypotheses and questions	<i>72</i>
1.6 Publications arising from data presented in this thesis.	<i>73</i>
Co-authorship statement	<i>74</i>
Chapter 2: Methods	<i>76</i>

2.1	B-ALL Wet lab methods	76
2.1.1	Immortalized Cell lines.....	77
2.1.2	Pediatric B-ALL primary lymphoblasts.....	83
2.1.3	Human biofluid samples.....	86
2.1.4	EV isolation optimization.....	87
2.1.5	EV characterization.....	89
2.1.6	RNA isolation and processing.....	92
2.2	B-ALL dry lab methods	94
2.2.1	miRNAs in pediatric B-ALL cells.....	94
2.2.2	Pediatric B-ALL EV Small RNA.....	100
Chapter 3: Meta-analysis of microRNA profiling data does not reveal a consensus signature for B cell acute lymphoblastic leukemia.		104
3.1	Introduction	104
3.2	Results	105
3.2.1	Comparison of healthy versus disease miRNA profiles revealed miRNA expression patterns.....	105
3.2.2	Meta-analysis of DEmiRs show conflicting expression patterns for some miRNAs.....	108
3.2.3	Target prediction identifies mRNA targets.....	114
3.3	Discussion	121
3.4	Conclusion	127
Chapter 4: Extracellular vesicle small RNA cargo discriminates non-cancer donors from pediatric B cell acute lymphoblastic leukemia patients.		128
4.1	Introduction	128
4.2	Results	131
4.2.1	Pediatric B-ALL plasma contains more extracellular vesicles than NCD.....	131
4.2.2	Vn96 isolates a more comprehensive EV population than SEC or ExoQuick.....	133
4.2.3	NCD and pediatric B-ALL EVs package different types of small RNA.....	134
4.2.4	A miRNA-specific signature does not clearly distinguish NCD from pediatric B-ALL.....	135
4.2.5	Sample heterogeneity of plasma EVs.....	135
4.2.6	EV RNA transcripts discriminate pediatric B-ALL from NCD.....	135
4.2.7	Both fragments and whole transcripts are packaged into EVs.....	146
4.2.8	Gene set enrichment analysis (GSEA) shows that EVs package gene sets that negatively regulate the cell cycle.....	146
4.2.9	GO and KEGG pathways of exclusive signature.....	151
4.2.10	Competitive endogenous RNA (ceRNA) analysis shows miRNA: mRNA pairs.....	151
4.3	Discussion	153
4.4	Conclusion	165
Chapter 5: Small RNA transcripts are selectively packaged into pediatric acute lymphoblastic leukemia-derived extracellular vesicles.		166
5.1	Introduction	166
5.2	Results	170

5.2.1 Pediatric B-ALL lymphoblast and cell line EVs have similar size and concentration.	170
5.2.2 Different types of RNA transcripts are packaged into EVs.	170
5.2.3 RNA transcript packaging mechanisms.	174
packaged into EVs, in contrast to other RNA types shown previously, which appear to be fragments.	178
5.2.4 Non-selectively packaged RNAs are enriched in cellular proliferation and metabolism.	178
5.2.5 Cell exclusive transcripts showed cell maintenance associated gene sets and pathways while EV exclusive transcripts include pediatric B-ALL and immune associated genes.	178
5.2.6 Sequence motifs are shared within cell exclusive and EV exclusive transcripts.	186
5.2.7 Different lncRNA and tRNA fragment subtypes are packaged into EVs.	191
5.3 Discussions	191
5.4 Conclusion	198
<i>Chapter 6: The effect of Streck preservative treatment of blood collection tubes on EV collection</i>	200
6.1 Introduction	200
6.2 Results	201
6.2.1 Samples collected in EDTA and Streck cell-free BCT.	201
6.2.2 Comparison of blood plasma and urine EV morphology between collection tubes.....	205
6.2.3 Comparison of differences in EV RNA cargo between collection tubes.....	211
6.2.4 Comparison of specific EV RNA content	213
6.3 Discussion	217
6.4 Conclusion	224
<i>Chapter 7: Discussion, next directions, and significance</i>	225
7.1 Discussion	225
7.2 Future directions	229
7.2.1 Exploring EVs as biomarkers for monitoring pediatric B-ALL.....	229
7.2.2 Elucidating the mechanism of RNA packaging into EVs.....	231
7.3 Significance	232
<i>Chapter 8: References</i>	235
<i>Appendices</i>	318
Appendix A: Copyright licenses	318
Appendix B: Codes and scripts for bioinformatics	321
B1: Ethics approval for pediatric B-ALL EV small RNA signature and packaging study	321
B2: Ethics approval for effect of storage conditions on EV isolation study.....	322
Appendix C: Codes and scripts for bioinformatics	323
C1: Codes and scripts for chapter 3	323
C2: codes and scripts for chapter 4.....	331
C3: codes and scripts for chapter 5.....	340
<i>Supplemental files and figures</i>	346
Supplementary figures and files for miRNA signature of pediatric B-ALL (Chapter 3).....	346

Supplementary figures for EV-based RNA signature of pediatric B-ALL (Chapter 4)	359
Supplementary figures for RNA sorting mechanisms in pediatric B-ALL (chapter 5)	363

Acknowledgement

'It takes a village to raise a child, " says a popular African proverb. In this context, my Ph D. truly took a village to get through, so I want to acknowledge the village that helped me along.

This could not have been possible without *YESHUA HAMASHIAH*, who keeps helping me every step of the way. I also want to thank my mother, Ibiye – whose sacrifices keep me looking forward.

I am also grateful to Ibis Longjohn – my brother and one of my best friends; and my sister Queeneth – who keeps me grounded. I am also grateful to my father, who inspires me in his own way.

My gratitude also goes to Mrs. Daba Soberekon and Mrs. Queen, without whom my educational journey would not have been possible. I also thank people who believed and continue to believe in me - Mr. Abiye Ford Briggs, Chief Dr. A. Okiri, Mrs. Aquilia Wokoma-Anji and her family, and Engr. D. Dr. Shreyasi and Shramana Sarkar, Sarika Kumari and Hong-Dien Anthony Phan were also great supports. I am grateful for my late Aunts (Abigail Raymond, Constance Soberetonari and Love Ihunwo), who believed in me but didn't live long enough to see me reach this point.

I want to thank Dr. Sherri Christian for saying yes to that ambitious girl who emailed her keenly. You welcomed me into your lab as an ambitious dreamer, and now I leave your lab as a confident (more accurately - trying to be) scientist who is still very ambitious. Wherever I go in my career, I will always remember your excellent supervision, kindness, support, and guidance. I am also grateful to Dr. Aimee Surprenant for all her help and support. Finally, I am grateful to my committee members - Dr. Andrew Lang, and Dr. Mark Berry, who guided my trajectory. I also thank the Beatrice Hunter Cancer Research Institute for their funding, Dr. Jo-Anna Hudson, and

Dr. Paul Moorehead. I am also grateful to the BC Children's Hospital Biobank, Vancouver, BC, and the Janeway Hospital St. John's for the samples that allowed me to do my research.

List of Tables

- Table 1.1. Types of antibodies idiotypes based on IgH chain rearrangements.
- Table 1.2. B cell subsets are classified based on their cell surface marker expression.
- Table 1.3. B-ALL subtypes using the Basel nomenclature and chromosomal abnormalities.
- Table 1.4. Extracellular vesicle physicochemical properties and biogenesis pathway.
- Table 2.1. Cell line study models used for the study.
- Table 2.2. Summary B cell immunophenotype of cell line study models.
- Table 2.3. Characteristics of patient or non-cancer donor (NCD) blood plasma samples
- Table 4.1. Biological characteristics for samples.
- Table 4.2. Total mature and precursor miRNA cargo of NCD and pediatric B-ALL EVs.
- Table 5.1. Biological characteristics for samples.
- Table 5.2. Read counts and small RNA output of cell and EV RNASeq.
- Table 5.3. Summary of number of detected pre-miRNAs using sRNADE platform.
- Table 6.1 Sample details for Streck/ non-Streck study blood collection.
- Table 6.2 Sample details for Streck/ non-Streck study urine collection.
- Table 6.3 Primer efficiencies and R^2 for plasma EV miRNAs.
- Table 6.4 qRT-PCR Ct values for three miRNAs from the one plasma sample pair.
- Table 6.5 Top 5 selected miRNAs for urine EV qPCR.
- Table 6.6 Primer efficiencies and R^2 for urine EV miRNAs.

List of Figures

- Fig. 1.1 Schematic illustration of cellular and non-cellular components of human peripheral blood and Hematopoiesis in the bone marrow.
- Fig. 1.2. A typical structure of an immunoglobulin (antibody).
- Fig. 1.3. B cell development is a discretely regulated process.
- Fig. 1.4. The effector functions of antibodies.
- Fig. 1.5. Types of ribonucleic acids (RNA).
- Fig. 1.6 The estimated incidence of four major groups of hematological malignancies in Canada. Data obtained from the Leukemia and Lymphoma Society of Canada
- Fig. 1.7. Extracellular vesicles characteristics.
- Fig. 1.8. A summary of the bioactive cargo of EVs from hematological malignancies and their downstream effects upon EV uptake by recipient cells.
- Fig. 1.9. Comparison of tumour biopsy techniques.
- Fig. 1.10. Schematic diagram showing EV-based communication between B cell lymphocytes, B-ALL lymphoblasts, and stromal cells.
- Fig. 2.1. Summary of wet lab experiments.
- Fig. 2.2. Cell microphotography of cell line study models.
- Fig. 2.3. Sample flow cytometry data showing expression of key surface and intracellular markers of B cells of different cell lines.
- Fig. 2.4.** Experimental design for small RNASeq analysis to identify an EV signature.

- Fig 3.1 Experimental design.
- Fig. 3.2. Analyses of differentially expressed miRNAs (DEmiRs) reanalysis.
- Fig. 3.3 Re-analysis of raw profiling data identified overlapping miRNAs and DEmiRs (between datasets) but no clear consensus signature.
- Fig. 3.4. Meta-analysis of differentially expressed miRNAs (DEmiRs) from published papers show a lack of consistency of DEmiRs across multiple platforms.
- Fig. 3.5. Network analyses of biological processes and gene pathways.
- Fig. 4.1 Experimental design for small RNASeq analysis to identify a signature.
- Fig. 4.2. Characterization of EVs in non-cancer donors (NCD) and pediatric B-ALL plasma.
- Fig. 4.3. A miRNA-specific EV signature does not distinguish pediatric B-ALL from non-cancer donors.
- Fig. 4.4. RNASeq data exploration.
- Fig. 4.5. Analysis of small RNA using DESeq2 distinguishes pediatric B-ALL from NCD.
- Fig. 4.6. Analysis of small RNA exclusive to B-ALL and NCD is highly discriminatory.
- Fig. 4.7. Visualization of RNA-seq data from selected DPRNAs using an integrative genome viewer (IGV) showed different RNA packaging patterns.
- Fig. 4.8. Gene set enrichment analysis (GSEA) shows that transcripts that negatively regulate the cell cycle are enriched in B-ALL EVs

- Fig. 4.9. Functions enriched in NCD represent multiple cell types while functions in B-ALL are related to cancer metabolism.
- Fig. 4.10. Venn diagram showing mRNA targets of differentially packaged miRNA and mRNA cargo of EVs.
- Fig. 5.1. Typical characteristics of EVs from Vn96-isolation and within conditioned media.
- Fig. 5.2. Different RNA transcript sub-type abundance between EVs compared to the originating lymphoblasts or cell lines.
- Fig. 5.3. Data showing exclusively sorted RNAs are discriminating cells and EVs.
- Fig. 5.4. Fragments of introns and exons as well as full length RNAs are packaged into EVs and retained in cells.
- Fig. 5.5. Genes involved in cellular metabolism retained in cells while genes involved in immune function are packaged into EVs.
- Fig. 5.6. Gene set enrichment analyses (GSEA) shows that cells tend to retain general cellular maintenance and immune evasion gene sets while immune related and anti-proliferation gene sets are packaged in EVs.
- Fig 5.7. Cell exclusive RNAs share cell inclusionary motifs.
- Fig 5.8. EV exclusive RNAs share EV packaging motifs.
- Fig 5.9. RNA biotype specific motifs for significant RNA biotypes packaged into EVs.
- Fig 5.10. Different types of lncRNA are packaged in EVs and tRNA are differentially packaged depending on the amino acid.

- Fig 6.1 Quantification of urine EVs collected in both collection tubes. The top row is the EVs from the EDTA tube, while the bottom row is Streck BCT collected EVs.
- Fig 6.2 Quantification of EVs in blood collected in both sample collection conditions.
- Fig 6.3 Quantification of EVs in urine samples collected in non-Streck versus Streck tubes.
- Fig 6.4 Sample mirVana output for plasma EVs, showing an electropherogram of EDTA container collected plasma EV RNA
- Fig 6.5 Comparison of RNA in plasma EVs
- Fig 6.6 Comparison of the RNA in urine EVs
- Fig 6.7 qPCR primer validation for plasma miRNAs.
- Fig 6.8 qPCR primer validation for urine miRNAs.
- Fig 6.9 Relative levels of urine miRNAs

List of Abbreviation and Symbols

Abbreviation	Full meaning
AS	Antisense transcript
BCT	Blood collection tube
BM	Bone marrow
CLP	Common lymphoid progenitor
CSF	Cerebrospinal fluid
DC	Dendritic cell
DNA	Deoxyribonucleic acid
DT	Divergent transcript
EDTA	Ethylenediaminetetraacetic acid
EFS	Event free survival
EVs	Extracellular vesicles
FCM	Flow cytometry
HG	miRNA host gene
HSC	Hematopoietic stem cell
LINC	Long intergenic non-coding RNA
lncRNA	Long non-coding RNA

miRNA	Micro RNA
mRNA	Messenger RNA
MSC	Mesenchymal stromal/stem cell
MZ B cells	Marginal zone B cells
NTA	Nanoparticle tracking analysis
OS	Overall survival
PBS	Phosphate buffered saline
PEG	Poly-ethylene glycol
RBC	Red blood cells
RFS	Relapse free survival
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SEC	Size exclusion chromatography
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
TBS	Tris buffered saline
TEM	Transmission electron microscopy

TI antigen	T cell independent antigen
tRNA	Transfer RNA
Vn96	Venceremin 96
WB	Western blot
WBC	White blood cells

List of Appendices

Appendix A: Copyright licenses

Appendix B: Ethics approvals

Appendix C: Codes and scripts for bioinformatics

List of Supplemental Files and Figures

Supplementary files are attached to the PDF and can be accessed using the paperclip icon in Adobe Reader.

Supplementary files and figures for miRNA signature of pediatric B-ALL (Chapter 3)

Supp. fig. 3.1 The raw data of healthy controls (HC) versus B-ALL were normalized and the normalized data were shown in boxplots.

Supp. fig. 3.2 Average linkage hierarchical clusters for meta-analysis of scientific articles arranged based on type of analysis.

Supp. file 3.1 Excel file showing GEO ID, PMID, type of RNA analysis, miRNA analysis platform, B-ALL subtypes (when present), tissue source of samples, number of controls and patients and normalization techniques for reanalysis cohort datasets.

Supp. file 3.2 Promising DEmiRs and overlapping targets for miRDB and miRwalk databases.

Supp. file 3.3 List of miRwalk and miRDB predicted targets for 8 promising DEmiRs.

Supplementary files and figures for EV-based RNA signature of pediatric B-ALL (Chapter 4)

Supp. fig. 4.1 Sample NTA data for NCD and pediatric B-ALL samples, showing histogram – typical output of NS300 run.

Supp. fig. 4.2 Western blot image comparing EV isolation techniques.

Supp. fig. 4.3 Transmission electron microscopy (TEM) images of pediatric B-ALL cell line EVs from CCRF-SB cell line (top) and UoC-B1 (bottom) isolated with Vn96.

Supp. fig. 4.4 Unsupervised hierarchical cluster of small RNA EV clusters that are not clearly discriminatory.

Supp. file 4.1. List of miRNAs and expected packaged mRNA targets.

Supp File 4.2. List of EV RNAs exclusive to all 8 B-ALL or 6 NCD samples.

Supplementary files and figures for RNA sorting mechanisms in pediatric B-ALL (chapter 5)

Supp. fig. 5.1. RNA transcript types in cell and EV samples.

Supp. fig. 5.2. Gene ontology (GO) and Kyoto encyclopedia of gene and genomes (KEGG) pathways for top 200 non-selectively packaged RNAs.

Supp. file 5.1. Cell and EV exclusive RNAs ranked based on average expression levels.

Supp. file 5.2. List of sample tRNAs and lncRNAs packaged from cells to EVs.

Supp. file 5.3. List of motifs common to all cell exclusive RNAs.

Supp. file 5.4. List of motifs common to all EV exclusive RNAs

Chapter 1: Introduction

1.1 B Lymphocytes

The immune system is a system that has evolved to protect an organism against infection. The immune system is composed of a complicated and dynamic network of cells, organs, molecules, and pathways¹ (**Fig. 1.1A**). Lymphocytes are immune cells that offer protection in a pathogen specific manner (adaptive immune system). Lymphocytes fall into three categories, with each distinct in their characteristics, yet interwoven in their mode of action. These are B lymphocytes (B cells), which are bone marrow (BM) associated and produce antibodies; T lymphocytes (T cells) which are thymus associated, support other immune cells and eliminate transformed self-cells; and natural killer cells (NK cells) which eliminate malignant and infected cells². Together or independently, these cells participate in eliciting appropriate immune responses. This thesis and the associated research will focus largely on B lymphocytes, covering aspects of normal B cell development, and delving further into a disease that arise because of aberrant B cell development.

1.1.1 Discovery of B cells

The discovery of antibodies predated that of B cells. To identify which cells produce antibodies, two teams of immunologists across the world worked simultaneously in search of the source of antibodies. While one team looked at the chemistry of lymphocytes, the other team looked at the molecular biology of lymphocytes³. Notably, in 1956, Max Cooper at Robert Good's laboratory at the University of Minnesota, Minneapolis discovered that irradiated chickens whose Bursa of Fabricius had been surgically removed produced no antibodies when injected with bovine serum albumin or the bacterium *Brucella abortus*⁴. In 1957, the clonal-selection theory was developed

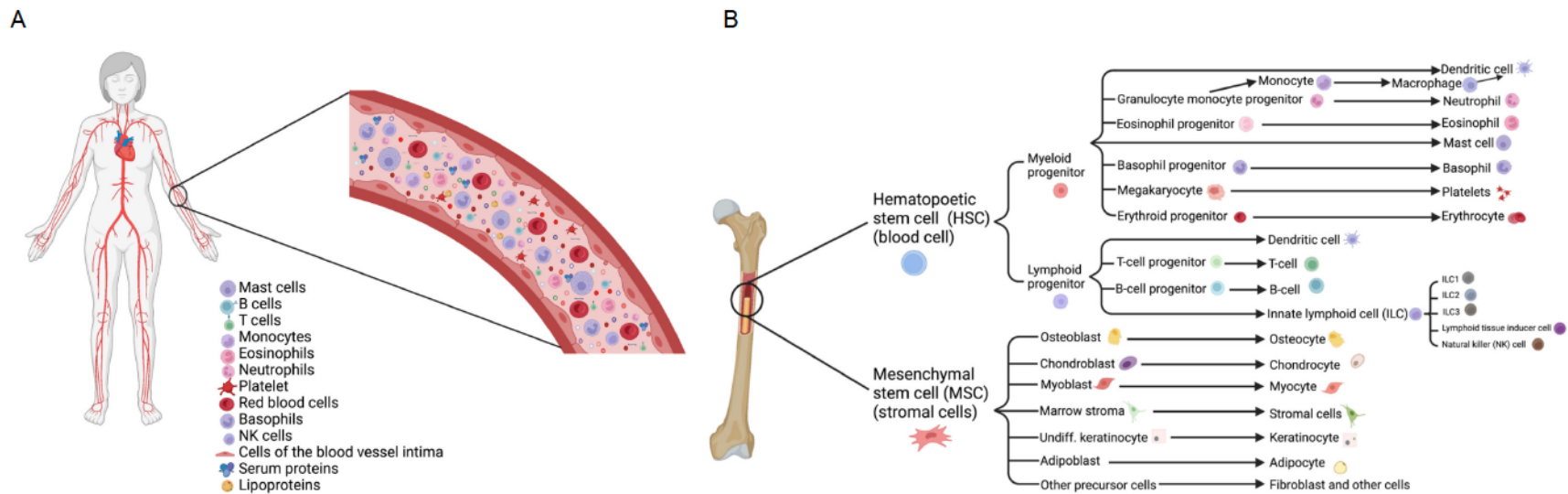


Fig. 1.1. (A) Schematic illustration of cellular and non-cellular components of human peripheral blood. Peripheral blood involves vessels that carry or allow the flow of blood systemically – to and from specific organs. The walls of the blood vessel have three layers, the intima, media, and adventitia, which function together to maintain blood, enabling the delivery of nutrients, waste, and other substances between compartments in the body. The cells present in blood include Leukocytes (white blood cells - WBC), erythrocytes (red blood cells - RBC) and megakaryocytes (platelets). WBCs are immune cells and include agranulocytes (B and T lymphocytes, natural killer (NK) cell and monocytes), and granulocytes (eosinophils, neutrophils, basophils, and mast cells). Each blood immune cell carries out specific functions and works in tandem with other non-immune blood cells when necessary. **(B)** Within the bone marrow (BM), there are two sets of cells, based on the location within the bone marrow, the expression of specific markers and the potential to form specific cells. Hematopoietic stem cells (HSCs) give rise to blood cells – myeloid and lymphoid progenitors. The progenitors of

the HSC lineages give rise to blood cells - leukocytes, erythrocytes, and megakaryocytes. Mesenchymal cells are stromal cells, that helps the maintenance of BM environment and the blood cell's growth and integrity. The BM stromal cell progenitors have homing mechanisms, based on expression of chemotactic factors⁵ that home cells to different locations where the cells differentiate into specialized cell types.

by Sir Frank Macfarlane Burnet at the Walter and Eliza Hall Institute (WEHI) of Medical Research in Melbourne and David Talmage at the University of Chicago, Illinois^{6,7}. Based off the 1956 discovery and in the light of the 1957 proposed theory, in 1974, Max Cooper, Martin Raff and John Owen at the University College London and a team of Gustav Nossal at the WEHI and Pierre Vassalli at the University of Geneva, Switzerland found that in mouse, hematopoietic (blood-forming) tissues in mammals are the equivalent of the Bursa^{3,8,9}.

1.1.2 Hematopoiesis

The network of cells that make up the immune system - blood cells in circulation interact in specific ways. Blood cell growth and development, called hematopoiesis, happens in preparation for maturation at different sites within the body and is critical to immune function. Blood cells develop, mature, and interact at primary, secondary and tertiary lymphoid organs. At primary lymphoid organs - the BM and thymus, hematopoietic stem cells (HSC) (**Fig. 1.1B**), including myeloid and lymphoid cells, grow and develop from immature precursors to differentiated cells in an antigen independent manner. Secondary lymphoid organs – spleen, lymph node (LN) and specialized lymphoid tissues, are sites where lymphocytes encounter cognate antigens, to stimulate differentiation into effector and memory cells. Secondary lymphoid organs regulate the expansion of lymphocytes in an antigen-dependent fashion. Finally, only in response to chronic inflammation or tissue injury, tertiary lymphoid tissues– collections of immune cells that mimic secondary lymphoid organs form within peripheral non-lymphoid tissues¹⁰⁻¹². Primary and secondary lymphoid organs are linked by blood and lymphatic vessels, allowing them to function together.

1.1.3 Normal B cell development

In the fetus, B cell development occurs in the liver, and moves to the BM post-birth. Therefore, B lymphocytes start and complete >75% of their growth and differentiation in the BM, with pre-B cells present in the endosteal niche associating with osteoblasts. The goal of this development is the generation of B cells which can express membrane-bound B cell receptor (BCR), which can specifically recognize and bind antigens, leading to the secretion of circulatory immunoglobulins. Immunoglobins (also called antibodies) are glycoproteins produced by plasma B cells, which are critical components of the immune response via recognition and binding to specific antigens (**Fig. 1.2**). B cell development starts as HSC and culminates as mature B cells, a process that takes about 1 – 2 weeks¹. The stages of B cell development are from common lymphoid progenitors (CLP), to progenitor (Pro-) to precursor (Pre-) to immature B cells, also called Hardy fractions (Fractions A -F) classified based on expression of different cell surface, membrane and nuclear markers (Basel nomenclature)¹³⁻¹⁵ (**Fig. 1.3**). These markers DNA methylases, transcription factors (TF) and enzymes, which are discretely regulated through the action of factors including noncoding RNA. B cell development occurs in two phases. The first phase is antigen-independent development – development that does not require the presence of a cognate antigen. Antigen-independent B cell development begins when HSCs goes through a serial continuum of progressive asymmetric cell divisions regulated by methyl transferases *DNMT1*, *DNMT3a* and *DNMT3b*¹⁶, and TFs *PU.1* and *Ikaros*¹⁷, which drive HSCs down the CLP lineage path. From the CLP, DNA methylases *MYSM1*, *CBP*, *HMT*, *HDAC3*, *SWI/SNF* and *Mi-2/NuRD* and TFs *E2A*, *EBF*, *Foxo1*, *IRF4/8*, *C-Myb* and *Runx1*^{1,18,19} facilitates B cell progression to the pre- and pro- stages, and subsequently culminating in Immature B cells. Immature B cells then progress through a transitional state when they undergo

Typical immunoglobulin Structure

CD79A (IgA), and CD79B (IgB)

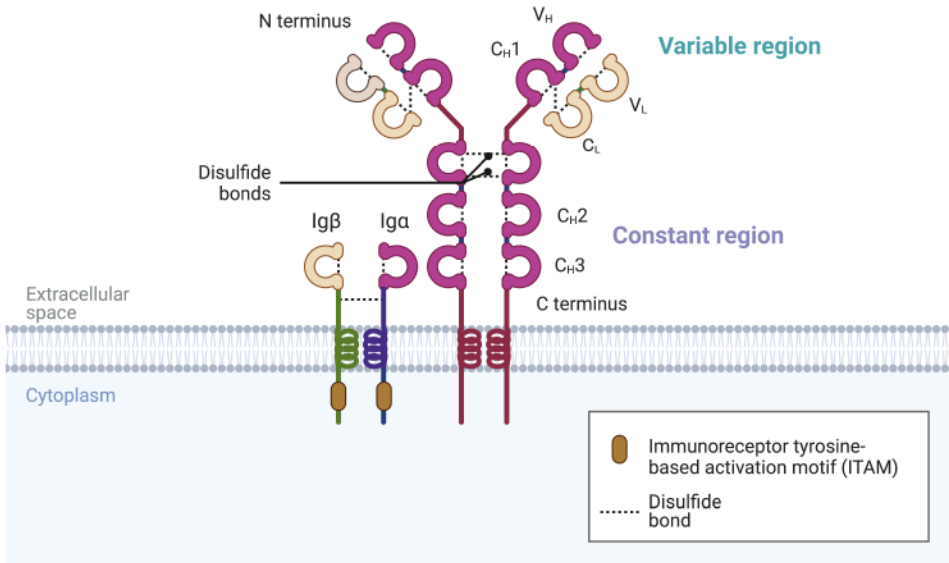


Fig. 1.2. A typical structure of an immunoglobulin (antibody). A typical antibody such as IgM is a polymer that consists of two long heavy chain molecules (purple) encased by two shorter light chains (cream). A heavy chain consists of an upper variable region (V_H), and a lower constant region consisting of three subunits (C_{H1} , C_{H2} and C_{H3}). The light chain contains one heavy and one constant region respectively. The heavy and light chain variable regions make up the hypervariable region – the portion of the antibody that is bound by the antigen. The heavy chains carry cytoplasmic tails which together with the CD79A/CD79B molecules facilitate intracellular signaling.

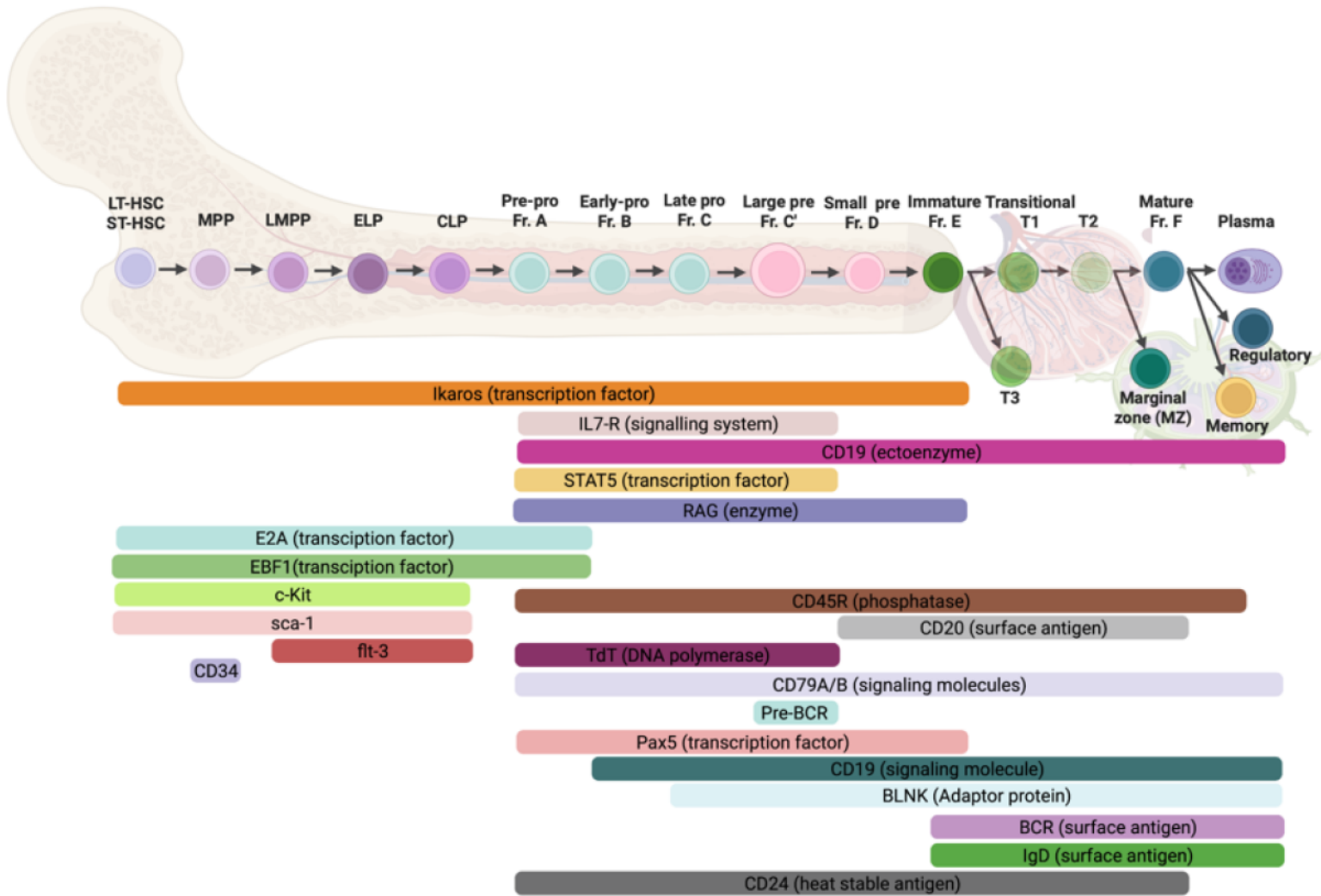


Fig. 1.3. B cell development is a discretely regulated process. In the bone marrow, a rare quiescent cell population called long-term hematopoietic cells (LT-HSC) differentiate into short-term HSC (ST-HSC) and subsequently into multipotent progenitors (MPP) that have no self-renewal ability. MPPs then differentiate into lymphoid-primed multipotent progenitors (LMPP) and subsequently into early lymphoid progenitors (ELP) that in turn give rise to common lymphoid progenitors (CLP). Antigen independent B cell development in the bone marrow starts when the common lymphoid progenitors (CLP) proceed to the next stage by an interplay of expression of cell surface, cytoplasmic and nuclear markers including signaling molecules, phosphatases, and adaptor proteins whose gene expression are in turn regulated by presence or absence of transcription factors. Thus, by the discrete interplay of these molecules, CLPs progress to pre-pro and early and late progenitor (pro) B cells. The marker expression landscape changes, with the presence/absence of these markers acting as pseudo checkpoints for progression of B cell development to the next stage. At the end of the large pro-B cell stage comes the Pre-B cell stage, where a pre-B cell receptor (pre-BCR) is expressed, leading to maturation into immature B cells capable of signaling through the BCR. Immature B cells can then migrate to secondary lymphoid organs where they encounter their cognate antigen to complete their antigen dependent maturation.

organs (lymph node and the spleen) where they continue development upon encountering a cognate antigen. Fully mature B cells can undergo clonal selection upon binding a cognate antigen and give rise to terminally differentiated antibody producing plasma B cells in circulation.

The second phase of B cell development is the antigen-dependent stage, that allows transitional B cells to progress to a more mature state. This phase involves secondary lymphoid organ- based regulation that culminates in the production of memory, effector, and antibody-producing plasma B cells^{10,20-22}. Antigen dependent development hinges solely on the availability and interaction of BCR with cognate antigen, which in turn leads to clonal selection and expansion.

Completion of the B cell development process results in mature, immunocompetent B cells, that express membrane-bound immunoglobulin that can specifically bind to antigens. Each individual B cell expresses approximately $1.5-3 \times 10^5$ copies of antibodies of unique specificity, with identical binding sites for antigen¹. In addition, through somatic hypermutation (SHM), B cells can improve their ability to bind antigens. SHM provides the molecular basis for affinity maturation, a key to optimizing antibody dependent immune responses²³. SHM involves point mutations accumulating in the V-regions of antibody's heavy and light chains, at a rate of about 10^6 orders higher than the background mutation in other genes²⁴. The goal of SHM is production of high-affinity antibodies; however, without selection SHM cannot decipher between favorable and unfavorable mutations, which could give rise to antibodies with either high, low or unchanged affinity²⁵. SHM is regulated by the action of the enzyme activation induced deaminase (AID)²⁶.

Upon antigen binding, activation of plasma cells result in antibody secretion, which has functions such as neutralization of infectivity, phagocytosis, antibody-dependent cellular cytotoxicity

Table 1.1. Types of antibodies isotypes based on IgH chain rearrangements.

	IgA	IgD	IgE	IgG	IgM
Heavy chains	α	δ	ϵ	γ	μ
	Monomer, Dimer	Monomer	Monomer	Monomer	Monomer, pentamer
	IgA1, IgA2	IgD	IgE	IgG1, IgG2, IgG3, IgG4	IgM
Number of antigen binding sites	4	2	2	2	10
Molecular weight (Da)	385,000	180,000	200,000	150,000	900,000
Percentage of total antibody in serum	13%	1%	0.0025%	80%	6%
Crosses placenta	Yes	No	No	No	no
Fixes complement	No	No	No	Yes	Yes

Fc binds to	-	-	Mast cells and basophils	Phagocytes	-
Function	Secreted into mucus, tears, saliva, colostrum	B cell receptor	Allergy and antiparasitic activity	Main blood antibody of secondary responses, neutralizes toxins, opsonization	Main antibody of primary responses, best at fixing complement; the monomer form of IgM serves as the B cell receptor

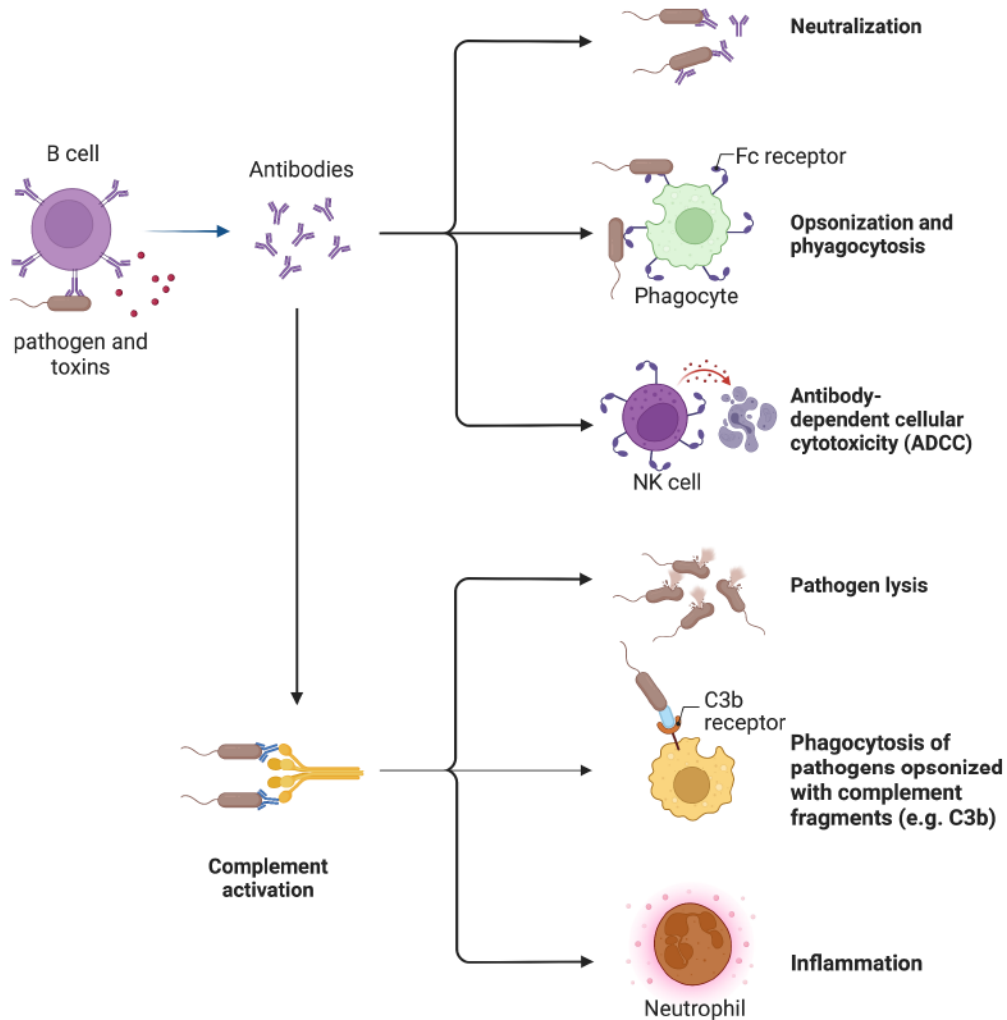


Fig. 1.4. The effector functions of antibodies. Upon successful antigen-BCR binding, clonal expansion leads to rapid proliferation and secretion of membrane antibodies. These antibodies in turn cause neutralization of the pathogens and toxins, opsonization, phagocytosis and cellular toxicity (in tandem with NK cells). The pathogens bound by the secreted antibodies are further eliminated by the complement system, which directly lysis the pathogens, phagocytosis of antibody opsonized pathogens and inflammation (in tandem with neutrophils).

(ADCC) and complement-mediated lysis of pathogen infected cells (**Fig. 1.4**). Each plasma B cell can secrete hundreds to thousands of antibodies per second¹. There are five classes of antibodies – IgM, IgG, IgA, IgE and IgD, which have different functions (**Table 1.1**).

Using class switching (which depends on SHM), antibodies of different functional classes can be generated. Activated B cells can also function as professional APCs (pAPC) – that internalize antigen through the antigen specific BCR. Activated B cells can also process and present antigenic peptides on their surface to T cells, while expressing costimulatory molecules (CD80 (B7-1) and CD86 (B7-2) that enable the activation of T cells²⁷. This interaction stimulates the T cells to release cytokines that induce the differentiation of the T cells into cytotoxic and helper T cells. T helper cells in turn activate macrophages to eliminate the pathogens and activate the B cells to differentiate into plasma and memory cells^{28,29}.

1.1.4 Types of B cells

Though the defining qualities of B cells are the presence of Ig receptors and the secretion of antibodies; not all B cells are the same. There are different types of B cells, which differ in their anatomical location, phenotype/cell surface marker expression (**Table 1.2**) and function¹. Specifically, B cell subtypes are classified based on marker expression phenotype, antigen specificity, developmental origins, and activation routes into B1 and B2 (follicular) populations, transitional, marginal zone, naïve (mature peripheral), plasma, memory, and regulatory cells.

Two main B cell lineages – B-1 and B-2 cells exist in the repertoire. B-1 B cells are more abundant in the peritoneal cavity than in the spleen³⁰. B-1 cells are abundant during early embryonic and early post-natal life and are subdivided into B-1a and B-1b B cells^{31,32}. B-1a cells account for at

Table 1.2. B cell subsets are classified based on their cell surface marker expression ^a.

B cell subpopulation	Basel Nomenclature	Marker expression
Hematopoietic stem cells	HSC	CD34 ⁺ CD117 ⁺ c-KIT ⁺ IL7R ⁺ ATXN1 ⁺ <i>FLT3</i> ⁺
Common lymphoid progenitor	CLP	CD34 ⁺ CD38 ⁺ CD117 ⁺ CD179B ⁺
Precursor progenitor B cell	Pre-pro-B	CD34 ⁺ CD38 ⁺ TdT ⁺ CD179A ⁺ CD179B ⁺ CD10 ⁺ IL7RA ⁺ CD19 ⁺ CD45 ⁺ CD43 ⁺ RAG1 ⁺ CD79A ⁺ CD79B ⁺
Early progenitor B cell	Early pro-B	CD34 ⁺ CD38 ⁺ CD24 ⁺ TdT ⁺ RAG2 ⁺ EBF1 ⁺ PAX5 ⁺ SOX4 ⁺ LEF1 ⁺ RAG1 ⁺ IGLL1 ⁺ CD179A ⁺ CD179B ⁺ CD79A ⁺ CD79B ⁺
Late Progenitor B cell	Late pro-B	CD10 ⁺ CD19 ⁺ IL7RA ⁺ CD19 ⁺ PAX5 ⁺ CD179A ⁺ CD179B ⁺ CD79A ⁺ CD79B ⁺
Large precursor B cell	Large pre-B	CD34 ⁺ CD38 ⁺ CD24 ⁺ IL7R ⁺ CD19 ⁺ CD45 ⁺ CD43 ⁺ RAG1 ⁺ CD179A ⁺ CD179B ⁺ RAG2 ⁺ IgK ⁻ IgL ⁻ IgH ⁺ CD79A ⁺ CD79B ⁺
Small precursor B cell	Small Pre-B	CD19 ⁺ IL7RA ⁺ IGH ⁺ RAG1 ⁺ RAG2 ⁺ SOX4 ⁺ LEF1 ⁺ EIF4EP1 ⁺ PAX5 ⁺ RSP27 ⁺ CD79B ⁺ CD79A ⁺ CD79B ⁺ CD24 ⁺

Immature B cell		CD19 ⁺ CD20 ⁺ CD24 ⁺ CD38 ⁺ CD45 ⁺ EBF1 ⁺ EIF4EBP1 ⁺ TNFRSF13C ⁺ IgK ⁺ or IgL ⁺ CD27 ⁻ IgM ⁺ IgD ⁻ CD10 ⁺ CD79A ⁺ CD79B ⁺
Mature B cell		IgM ⁺ IgD ⁺ CD20 ⁺ CD19 ⁺ EIF4EBP1 ⁺ TNFRSF13C ⁺ CD24 ⁺ IGHM ⁺ PAX5 ⁺ RSP27 ⁺ CD79A ⁺ CD79B ⁺
Plasma B cell		IgM ⁺ IgD ⁺ TNFRSF13C ⁺ IGKC/IGLC2 ⁺ CD20 ⁺ CD79A ⁺ CD79B ⁺
B-1a B cells	B-1a	IgM ^{hi} IgD ^{lo} CD23 ⁻ B220 ^{lo} CD19 ⁺ CD5 ⁺ CD79A ⁺ CD79B ⁺
B-1b cells	B-1b	IgM ^{hi} IgD ^{lo} CD23 ⁻ B220 ^{lo} CD19 ⁺ CD11b ⁺ CD5 ⁻ CD79A ⁺ CD79B ⁺
B-2 B cells	B-2	See below
Marginal zone B cells	MZ	CD19 ⁺ CD93 ⁻ CD21 ⁺ CD23 ⁻
Memory B cells	Double negative	IgD ⁻ CD27 ⁻ CD21 ⁺ CD24 ⁺ CD95 ⁺ CD19 ⁺ CD27 ⁺ CD38 ⁻ CD79A ⁺ CD79B ⁺
	IgM-only	IgM ⁺ IgD ⁻ CD27 ⁺ CD21 ⁺ CD24 ⁺ CD1c ⁺ CD79A ⁺ CD79B ⁺
Transitional 1 B cells	T1-B cells	CD19 ⁺ CD21 ⁻ CD23 ⁻ CD38 ⁺ CD24 ^{hi} CD10 ⁺ IgD ^{lo/-} CD93 ⁺ CD79A ⁺ CD79B ⁺

Transitional 2 B cells	T2-B cells	CD19 ⁺ CD93 ⁺ CD21 ⁺ CD23 ⁺ CD38 ⁺ CD24 ^{hi} CD10 ⁺ IgD ⁺ CD79A ⁺ CD79B ⁺
Transitional 3 B cells	T3-B cells	CD19 ⁺ CD93 ⁺ CD21 ⁺ CD23 ⁺ CD38 ⁺ CD24 ⁺ IgD ⁺ · CD79A ⁺ CD79B ⁺
Follicular	FO B cells	CD19 ⁺ CD93 ⁻ CD21 ⁺ CD23 ⁺ · CD79A ⁺ CD79B ⁺
Naïve B cells		IgD ⁺ CD27 ⁻ CD10 ⁻ CD79A ⁺ CD79B ⁺

^a Adapted from Peng et al 2018³³ and Kaminski et al, 2012³⁴

least 5% of total B cell populations in adult humans and self-replenish to maintain their numbers^{30,35,36}. They are self-reactive and persist in the body to produce natural antibodies that regulate immune reactivity³⁰. Conversely, B-1b cells are mostly located outside the secondary lymphoid organs and have a limited repertoire of antigen receptors³⁷. B-1 B cells secrete antibodies (primarily low-affinity IgM), which facilitates rapid response to T cell-independent antigens – antigens that do not require T helper cells for antibody production³⁸⁻⁴⁰.

B-2 B cells, are conventional immune competent B cells that exist in the body after discrete development⁴¹. All mature B cell types fall under this lineage of B cells. Naïve B cells have undergone antigen-independent development and are ready to leave the BM to continue the rest of the maturation process in a secondary lymphoid organ. Conversely, transitional B cells are temporal stages of B cells which have just left the BM and are brought to the spleen by blood. Transitional B cells are the bridge between BM immature and peripheral mature B cell stages⁴². These are divided into T1, T2 and T3 B cells, which are distinguishable by their cell surface receptor expression patterns⁴³ (**Table 1.2**). Within the spleen, T1 B cells enter the T cell zone, where they differentiate into T2 transitional B cells¹. Studies have shown that the order of progression of B cell differentiation is from T1 to T2 to mature B cells, a process that takes 3 – 4 days⁴⁴⁻⁴⁷. Generally, 75% of T1 B cells transition to T2 within the spleen while the remaining 25% emerge from the BM at the T2 state⁴⁸. The maturity level at T2 state is linked to changes in cytokine expression, which enables T2 B cells to recirculate within blood, and B cell follicles (follicular – FO cells) in the LN and spleen⁴⁹.

Within the follicles, T2 B cells undergo the final differentiation into mature T3 B-2 cells⁴⁹. A subsection of T2 B cells can enter the marginal zone and become marginal zone (MZ) B cells^{49,50}.

Contact of T3 B cells with soluble self-antigens within the spleen (or other organs) renders them anergic, without being eliminated from the B-cell repertoire^{51,52}. This interaction stimulates these cells, causing a failure to divide, differentiate or secrete antibodies and shortly after death⁵⁰.

Marginal Zone (MZ) B cells are resident in the spleen, with “innate-like” properties, being able to respond to T cell independent (TI) antigens⁵³. Because of their unique location in the spleen, they can respond rapidly to antigens carried by blood to the spleen. Antigens are presented to MZ B cells by MZ macrophages, neutrophils or DC in circulation⁵⁴. Though the MZ cells do not require T cell help to mount response to antigens, they receive help from myeloid cells in the form of secreted cytokines. MZ B cells are excellent at processing and presenting antigens. Finally, mature B cells are cells which have completed antigen-independent and-dependent development^{55,56}.

Memory B cells are differentiated B cells which are products of clonal expansion in response to antigen recognition by a B cell⁵⁷. Memory B cells remain in the bloodstream long after infection. Upon host re-exposure to the same infection, memory B cells can become quickly activated (with T cell help). Conversely Plasma B cells are effector cells, which are large B cells with very large endoplasmic reticulum, to enable the synthesis of large amounts of antigen specific antibodies. Plasma cells are CD38⁺ and are part of the two walls of memory (together with memory B cells)⁵⁷.

1.2 Gene expression for B cell development

Gene expression is the conversion of gene sequences (genomic DNA) into functional gene products such as proteins or functional RNAs. The human genome contains coding and non-coding genes, with coding genes accounting for only ~1.2% of the genome covering ~34Mb⁵⁸. As part of gene expression, regions of the genome that contain protein coding-genes are transcribed into mRNA (template for protein synthesis) ^{59,60}. Conversely, non-coding genes are transcribed into

non-coding RNA (ncRNA) that serve functional roles chromosome maintenance and chromatin remodeling and regulation of gene expression at the transcriptional and post transcriptional levels^{59,60}. Additionally, some ncRNAs are derived from intronic sequences of protein coding genes⁶⁰. Overall, a portion of the genome is transcribed into RNA, with the downstream process involving the DNA depending on the type of transcript (protein-coding versus non-protein coding). Additionally, gene expression is regulated pre- and post-transcriptionally by the action of general and specific transcription factors, and post translationally by the action of specific enzymes. Overall, gene expression is a highly regulated process, regulated at different layers by an interplay of factors including gene expression functional products.

B cell development is regulated by layers of interactions between different mediators at different levels. Firstly, transcription factors (TF) regulate the expression of development stage markers, making them the first key layer of regulation of B cell development. Expression of TF are in turn modulated by the action of non-coding RNAs such miRNAs and lncRNAs^{18,61}. Therefore, the second layer of regulation is non-coding RNA based modulation of TF, cytokines, and markers levels. miRNA and lncRNA modulation of B cell development TF could occur by direct binding or indirectly by altering cellular pathways that facilitate the expression of the TF¹⁸. Conversely, TFs upregulate the expression levels of miRNA and lncRNA to control cell fate¹⁸.

During B cell development, protein coding genes encode stage and lineage specific family of TFs such as PAX5, PU.1, E2A, and FOXO1 that regulate B cell development, by modulating B cell lineage commitment¹⁸. The main TF families includes the helix-turn-helix (e.g., Oct-1), helix-loop-helix (e.g., E2A), and the zinc finger (e.g., GATA proteins and glucocorticoid receptors). Other TF families are the basic protein-leucine zipper (e.g. cyclic AMP response element binding factor (CREB) and activator protein-1 (AP-1)) and β sheet motifs (e.g. nuclear factor KB

(NFKB)⁶². Other coding DNA encoded regulators of B cell development include general TFs – such as RNA polymerase II, promoter-specific activators and coactivators, and steroid hormones⁶³. B cell development is also regulated via chromatin remodeling, and posttranscriptional modification^{63,64}. After transcription of coding genes into mRNA, the initial product is premature mRNA (premRNA) transcript, which is then processed into the mature mRNA transcript. Mature mRNA then serves as a template processed by the ribosomal machinery into protein products.

Non-coding RNAs (ncRNA) are divided based on function into structural and regulatory (**Fig. 1.5**) respectively. Specifically, ncRNAs regulates epigenetic processes, including heterochromatin formation, histone modification, and gene silencing. Furthermore, ncRNAs regulate cellular function, leading to roles in health and disease. Ribosomal RNA (rRNA) is encoded by ribosomal RNA gene repeats, a major component of ribosomes. The major human rRNA gene repeat family is called the ribosomal DNA (rDNA), with each repeat consisting of 18S, 5.8S and 28S rRNA^{65,66}. Specifically, there are 200 – 600 copies of the rDNA gene among tandem arrays on the short arms of chromosomes 13 - 15, and 21 - 22⁶⁷⁻⁷². For rRNA biogenesis, rDNA is transcribed in the nucleolus by RNA pol I (18S, 5.8S and 28S) and III (5S)⁷³⁻⁷⁸. The rRNA molecules work in different combinations to form the ribosome: two subunits – the large subunit (LSU) consists of 28S, 5.8S and 5S while the 18S molecule makes up the small subunit (SSU)⁷⁹. rRNA functions in ribosome biogenesis, which regulates genome stability and cell cycle control⁸⁰⁻⁹¹. Transfer RNAs (tRNA) are 70 – 100 nucleotide (nt) long RNA, which have a cloverleaf secondary structure formed from Watson-Crick base pairing and creates helical stems that end in loops formed by unpaired bases⁹². tRNAs act as vehicles, bringing the amino acids needed to the growing polypeptide chain on the ribosome of the rough endoplasmic reticulum⁹³. In eukaryotes, tRNAs

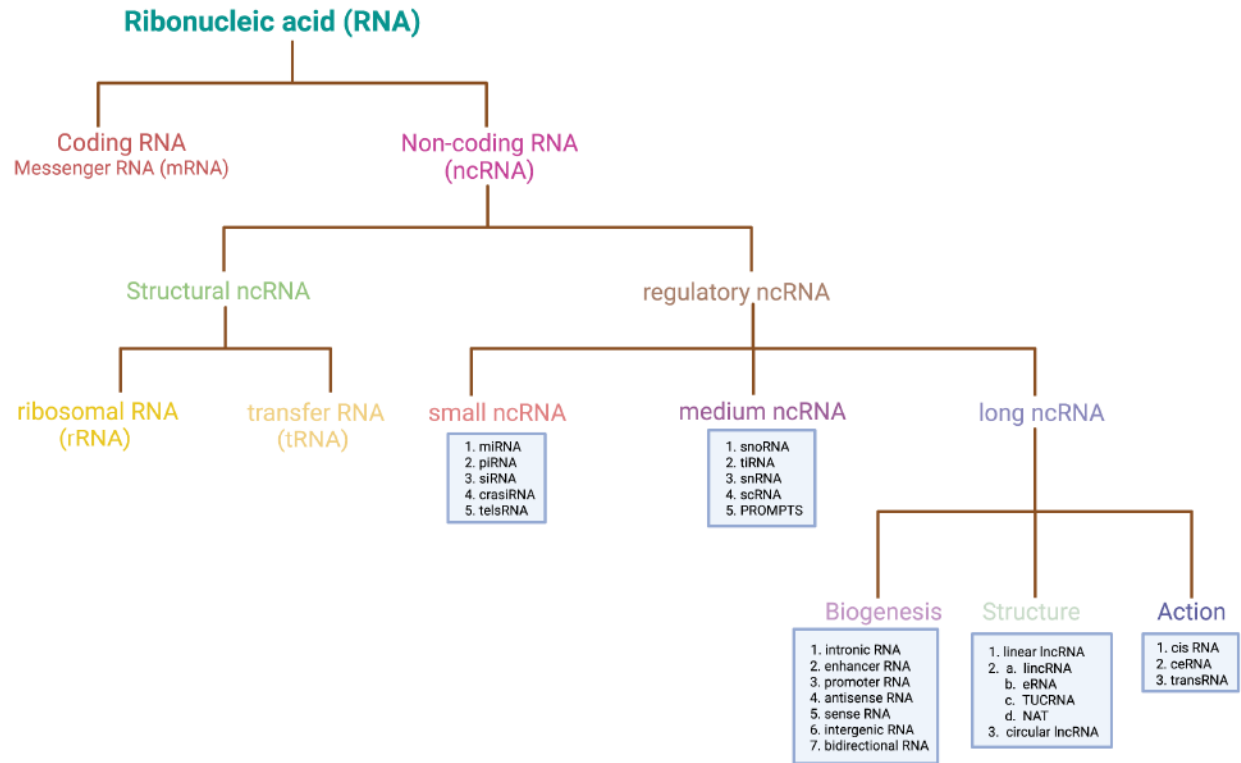


Fig. 1.5. Types of ribonucleic acids (RNA). Coding RNA is called messenger RNA (mRNA) and varies in size, which gets translated into proteins. Conversely, non-coding RNA regulate several aspects of gene expression.

are transcribed by RNA Pol III at 3×10^5 copies per cell from tRNA genes, into pre-tRNAs⁹⁴⁻⁹⁶. Pre-tRNAs are processed by RNases P and Z. tRNAs encoded by nuclear genomes are transcribed in the nucleus followed by export to the cytoplasm where they perform their functions⁹⁷. MicroRNAs (miRNA) are small ncRNAs of <25 nucleotides in length, that negatively regulate gene expression^{61,98}. miRNA biogenesis begins with the transcription of DNA sequences or from cluster genes which are only transcribed as miRNA molecules or together as polycistronic transcripts^{99,100}. Alternately, miRNAs may be synthesized from an intron or untranslated region (UTR) of a protein coding gene¹⁰⁰. miRNAs have two main biogenesis pathways – Drosha and Dicer dependent canonical and the Drosha and Dicer independent non-canonical pathways¹⁰¹. miRNA facilitate regulation of cellular processes including development, proliferation, and differentiation. Overall, miRNA regulates mRNA by binding (via miRNA response element – MRE), making a single miRNA able to target multiple mRNAs, and multiple miRNAs able to target individual mRNAs¹⁰²⁻¹⁰⁴.

Piwi interacting RNA (piRNA) are the largest class of ncRNAs, small RNAs of 24 – 31 nt which possess 2'-O-methyl modification sites at the 3' terminus and are abundantly expressed in mammalian testis and ovary^{105,106}. piRNAs are synthesized in a Dicer-independent manner from single-stranded precursor transcripts at intergenic regions (piRNA clusters) which regulate transposon activity¹⁰⁵. Transcription of dual-strand clusters into piRNA precursors depend on RNA Pol II, Rhino-Cutoff-Deadlock complex (RDC complex), Moonshiner, TATA-box binding protein related factor (TRF2) and 3' repair endonucleases (TREX)¹⁰⁷. piRNAs can pass on the memory of “self” and “non-self”, suggesting a role in transgenerational cellular processes¹⁰⁵. piRNA also regulates gene expression and are altered in diseases including cancer.

Centrosome-associated short interacting RNAs (crasiRNA) are ~34 – 42 nt and are derived from centrosomes^{108,109}. They have been shown to regulate local chromatin modifications.

Telomere-specific small RNA (tel-sRNA) are ~24 nt long, which are synthesized in a dicer-independent manner and have 2' methylation at their 3' terminus. They are asymmetrical with specificity toward telomere G-rich strand and is evolutionarily conserved from protozoa to mammal¹¹⁰. Their function though currently unknown is suggested to be in telomere maintenance.

Small nuclear RNAs (snRNA) are small, highly abundant nuclear localized ncRNAs of ~ 150 nt, which are important for RNA processing including splicing of introns from pre-mRNA¹¹¹. snRNAs are complexed with proteins to form RNA-protein complexes called snRNPs within the nucleus, which make up the spliceosome¹¹². snRNPs (U1, U2, U4, U5 and U6) in combination with RNA-RNA complexes work to recognize intronic splice sites and branch sites followed by complex RNA-RNA and RNA-protein interactions culminating in the assembly of catalytically competent spliceosomes, with the snRNAs and their bound proteins playing central roles¹¹³.

Small nucleolar RNAs (snoRNAs) are 60 – 300 nt in length, and mainly accumulate in the nucleoli^{114,115}. In vertebrates, many snoRNAs are encoded within introns of protein-coding or non-coding genes, while some are autonomously transcribed by RNA polymerase II¹¹⁶. For many intronic snoRNAs, host gene co-transcription, splicing, intron lariat debranching and exonucleolytic digestion in the nucleoplasm is also part of their biogenesis¹¹⁵. SnoRNAs mainly regulate processing of rRNA, mRNA splicing and editing, stress response and metabolic homeostasis, and maturation of rRNA, snRNAs and other type of RNAs¹¹⁷. SnoRNAs are also involved in cancer regulatory pathways such as p53, and in tumorigenesis, tumor aggressiveness and staging¹¹⁵.

tRNA-derived and stress-induced small RNAs (tiRNA), also known as tRNA halves are a novel class of small ncRNAs of 30 – 45 nt, which are derived from tRNAs cleaved under stress conditions^{118–120}. tiRNAs and smaller tRNA fragments (tRFs) of 14 - 30 nt make up a group of novel small ncRNAs called tRNA-derived small RNAs (tsRNAs); which are classified based on their length and cleavage sites^{121,122}. For the specific biogenesis of tsRNAs, stressors such as hypoxia, amino acid deficiency, UV radiation, heat shock, oxidative damage and viral infection induce cleavage of mature tRNAs by angiogenin (ANG) in the anticodon region, to generate 5' tiRNAs (30-45 nt) and 3' tiRNAs^{118,119,123,124}. tRFs are derived from mature tRNAs or pretRNAs¹²⁵. These fragments may be linked to biological functions in health and disease.

Promoter upstream transcripts (PROMPTS) are a newly discovered class of RNA with a heterogenous length distribution (>200 nt) synthesized upstream of promoters of active protein-coding genes. PROMPTS are mostly localized in the nucleus, followed by rapid turnover via exosome mediated degradation, leading to short half-lives^{126,127}. PROMPTS may regulate transcription, DNA methylation and enforcing promoter orientation of protein-coding genes¹²⁸.

Long non-coding RNA (lncRNA) are >200nt in length, and poorly conserved across species. lncRNAs are a large and highly heterogenous collection of transcripts that differ in their biogenesis and genomic origin. lncRNAs can be transcribed from enhancer sequences, their own promoters or from promoters shared with divergently transcribed coding or non-coding genes^{129,130}. Therefore, there are lncRNAs transcribed by RNA polymerase II and other RNA polymerases, lncRNAs from intergenic regions (lincRNAs), sense transcripts, antisense transcripts, and divergent transcripts^{131,132}. lncRNAs have distinct features that are unique to them, including having a generally lower expression, fewer exons, and greater tissue specificity. lncRNAs can also interact with other RNAs, DNA and proteins, and a combination of these molecules, making their

functional repertoire broad, including modulation of epigenetic, transcriptional, translational and post-transcriptional events^{133,134}. lncRNAs have differential localization within cells, even though a greater proportion of them are initially localized in the nucleus¹³⁵⁻¹³⁷.

Competitive endogenous RNA (ceRNA) are RNAs that compete and interact for miRNA response elements (MRE) sites, through which a regulatory network called ceRNA networks (ceRNET) is formed. By acting as sponges for miRNAs that share MREs, ceRNAs create a network that affects the normal ability of miRNAs to targeting mRNAs^{138,139}. Other transcripts that can act as ceRNA are pseudogenes, lncRNAs, and circular RNA¹⁴⁰⁻¹⁴⁶. The concentration of miRNAs especially in relation to mRNA targets is a determinant of ceRNA activity. When there is a lower concentration of miRNAs than targets, the targets are unreplaced, leading to low ceRNA activity^{141-143,147}. Conversely, in the presence of higher miRNAs than targets, there would be an almost universal target repression, thus higher ceRNA activity¹⁴⁰. ceRNAs function in regulating biological processes, in health and disease, including emerging data in cancer, including B cell cancers¹⁴⁸.

1.3 Aberrant B cell development and hematological malignancies (HM)

Hematopoiesis is so discretely regulated, that a loss of regulation gives rise to diseases such as cancer. Cancers of immune cells - leukemias (originate in the BM) and lymphomas (originate in lymph nodes (LN) and other tissues)¹⁴⁹ are collectively called hematological malignancies (HM). Leukemia and lymphoma make up ~ 67% of all HM diagnosed in Canada, with relative proportions globally¹⁵⁰⁻¹⁵². Less common are multiple myeloma (MM) and myeloproliferative disorders.

1.3.1 Leukemia

Leukemia is characterized by symptoms that are initially related to bone marrow dysfunction, such as infections resulting from low neutrophil count; and bruising and bleeding, caused by

thrombocytopenia. Once the BM is at capacity, leukemia cells can infiltrate the LN, causing lymphadenopathy, or enter the circulation, liver, and spleen¹⁵³. Based on which cell lineage is involved, leukemia is classified into myeloid or lymphoid (**Fig .1.1**). It is further sub-categorized based on the stage of the cell at transformation and the growth rate of the neoplastic cells. Leukemia originating from developing immune cells are generally fast growing (acute), while those from cells at more mature stages are slow growing (chronic)¹⁵³. These classifications give rise to the four main categories of leukemia: acute lymphoblastic leukemia (ALL), acute myelogenous leukemia (AML), chronic myeloid leukemia (CML), and chronic lymphocytic leukemia (CLL) (**Fig. 1.6**).

1.3.2 B lymphoblastic leukemia

Lymphoblastic leukemia (B-ALL) symptoms tend to be nonspecific, thus a multi-pronged approach is taken for ALL diagnosis. The diagnosis workup for ALL includes BM aspiration and review of biopsy material for hematopathology parameters including morphologic assessment of Wright-Giemsa-staining of BM aspirate smears, and hematoxylin and eosin staining of core biopsy and clot sections¹⁵⁴. In addition, flow cytometry (FC) based immunophenotyping use panels that include antibodies against CD10, CD19, CD20 and CD22 for B-ALL and CD1a, CD5, CD8 for T-ALL. Minimal/measurable residual disease (MRD) in the BM at diagnosis is measured using baseline flow cytometry, molecular characterization of leukemic clones and karyotyping of G-banded chromosomes at metaphase¹⁵⁴. Furthermore, molecular analysis for optimal risk stratification based on clinical and genetic risk factors including presence of specific recurrent genetic abnormalities inform low, standard, intermediate, or high-risk designation, which

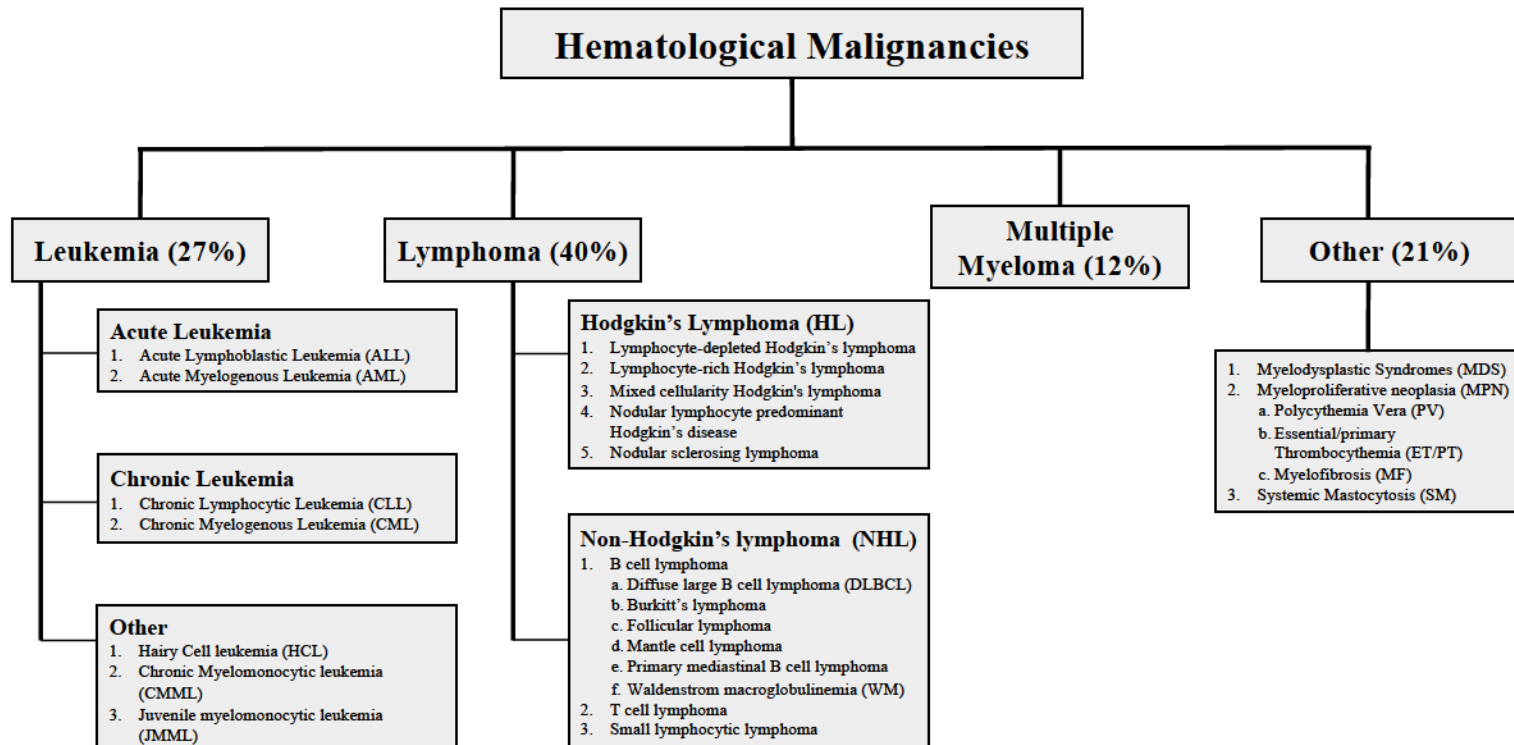


Fig. 1.6. The estimated incidence of four major groups of hematological malignancies in Canada. Data obtained from the Leukemia and Lymphoma Society of Canada¹⁵⁰.

in turn dictate favorable, intermediate, or poor outcome. Risk is determined based on age, white blood cell (WBC) count ($>30 \times 10^9/L$ for B-ALL and $>100 \times 10^9/L$ for T-ALL) at diagnosis and presence of specific rearrangements. A rearrangement such as ETV6::RUNX1 is associated with a favorable outcome while KMT2A is a poor prognosis associated rearrangement. Specifically, interphase fluorescence in situ hybridization testing, real time reverse transcriptase polymerase chain reaction testing, next generation sequencing and chromosomal microarray assays on peripheral blood and BM aspirates facilitate the identification of risk associated mutations¹⁵⁴. Central nervous system (CNS) involvement (CNS⁺) is diagnosed by lumbar puncture followed by microscopic examination of WBCs in cerebrospinal fluid (CSF) ^{155–157}. Testicular involvement at diagnoses and relapse can be diagnosed based on increased size, abnormal swelling and/or firmer consistency of the testis and eventually migrate into peripheral blood circulation where abnormal lymphoblasts crowd out other cells. Classification of B-ALL is based on the expression profile of B cell markers, such as CD10, CD19, CD20, CD22, CD34, HLA-DR, and CD79a, that are linked to the stage of B cell development when developmental arrest occurs¹⁵⁸. Thus, based on cellular markers, B-ALL is subtyped into pro-B-ALL, early pre-B-ALL, pre-B-ALL (common acute lymphoblastic leukemia [CALLA] and non-CALLA) and mature B-ALL.

Another basis of B-ALL classification is chromosomal abnormalities. There are two groups of B-ALL-associated genetic abnormalities, numerical and structural abnormalities. Numerical abnormalities (Hyperdiploidy and hypodiploidy) is characterized by a gain or loss of chromosomes; while structural abnormalities, such as TCF::PBX1, KMT2A rearrangements, and ETV6::RUNX1, involve translocations that result in chimeric protein formation or moving a gene close to a strong transcriptional promoter resulting in gene overexpression^{159–162} (**Table 1.3**).

Hyperdiploidy (>50 chromosomes) is the most prevalent mutation in pediatric B-ALL and is generally associated with a favorable outcome. However, high hyperdiploidy contains additional chromosomal gains and genetic rearrangements some of which could influence disease outcome including adverse outcome in about 25% of all hyperdiploidy patients^{159,163,164}. Specifically, duplications of 1q and 17q isochromosome, aneuploidy, non-random gain of specific chromosomes (such as +X, +4, +6, etc.), trisomies and tetrasomies are observed, based on cytogenetic analyses^{159,164}. Individual gains of chromosomes 4, 6, 10, 17 and 18 lead to better prognosis while isochromosome 17q and trisomy 5 lead to worse prognosis¹⁶⁵⁻¹⁶⁸. Individual gains of chromosomes 5 and 20 have poor prognosis for currently unknown reasons^{169,170}.

BCR::ABL1 (reciprocal translocation between 22q11.23 breakpoint cluster region- BCR and 9q34.1 Abelson tyrosine-protein kinase 1 – ABL1) fusion produces a derivative chromosome 22 called the Philadelphia (Ph) chromosome. The protein product is a chimeric cytoplasmic protein with *ABL1* and *BCR* domains fused; to mediate constitutive oligomerization without requiring activating signals^{159,171-173}. This promotes aberrant tyrosine kinase constitutive activity, inducing aberrant signaling via pathways such as JAK2/STAT5, MAP kinases and PI3K/Akt^{159,171-173}. Ph⁺ pediatric ALL presents at an older age with high WBC count, French American-British (FAB) morphology L2 and high incidence of CNS involvement¹⁶³. The age at diagnosis affects the prognosis, with one to nine-year olds faring better than adolescents and young adults (AYA)¹⁷¹⁻¹⁷³. BCR::ABL1-like is a subset of *BCR::ABL1*⁻ pediatric B-ALL where patients have a gene-expression profile containing mutations that deregulate cytokine receptor and tyrosine kinase signaling¹⁷⁴. Thus, this subtype is sensitive to tyrosine kinase inhibitors (TKI) treatment¹⁷⁵.

E2A::HLF translocation (t(17;19) (q22;p13) encodes a chimeric protein which likely activates transcription of HLF regulated genes, which may inhibit apoptosis via aberrant up-regulation of

Table 1.3. B-ALL subtypes using the Basel nomenclature and chromosomal abnormalities ^a

(A) Marker expression			
Type of B-ALL ^b	Immunophenotype		
pro-B-ALL	CD19 ⁺ CD34 ⁺ CD22 ⁺ TdT ⁺ cytoplasmic CD79a ⁺ CD10 ⁺		
early pre-B-ALL	TdT ⁺ CD19 ⁺ CD10 ⁻ CD22 ⁺		
pre-B-ALL (CALLA)	CD19 ⁺ CD10 ⁺ CD19 ⁺ HLA DR ⁺ , cytoplasmic IgM ⁺ CD22 ⁺ CD34 ⁺ TdT ⁺ cytoplasmic CD79a ⁺		
pre-B-ALL (non-CALLA)	CD10 ⁻ CD19 ⁺ HLA DR ⁺ , cytoplasmic IgM ⁺		
Mature B-ALL	CD10 ⁺ CD19 ⁺ CD20 ⁺ CD22 ⁺ surface IgM ⁺		
(B) Chromosomal abnormalities			
Established subtypes			
Subtype	Percentage (%)	Mutations	Prognosis
Hyperdiploidy	30	>50 chromosomes	Favorable
ETV6::RUNX1	25	t(12;21)(p13;q22)	Favorable
MLL rearranged (KMT2A)	9	11q23 rearrangement; t(4;11)	poor
TCF::PBX1 (E2A::PBX1)	6	t(1;19)(q23;p13)	Favorable

BCR::ABL1	3	t(9;22)(q34;q11.2)	Favorable
Hypodiploidy	1	<50 chromosomes	poor
Provisional subtypes			
BCR::ABL1 like (IKZF1, CRLF2 and JAK mutations)	8		poor
iAMP21	2	Mutations in chromosome 21	poor
New subtypes			
DUX4-rearranged (IGH, ERG)	4	Deletion of ERG	favorable
ETV6::RUNX1-like	3	t(17;19)(q22;p13)	Poor
ZNF384-rearranged	1	(p13.2;p11.2)	
MEF2D-rearranged	0.5	Mutation in chromosome 21	Poor
Currently unclassified/unestablished			
Others	10%		
c-MYC		t(8;14), t(2;8),t(8;22)	Favorable
IKAROS (IKZF1)			poor

a. *Table adapted from Chiaretti et al, 2014¹⁷⁹ Arber et al, 2016¹⁸⁰, Lilljebjorn et al, 2016¹⁷⁸, Lilljebjorn & Fioretos, 2017¹⁸¹ and Brown et al 2020¹⁸²*

b. *Classification of B-ALL subtypes based of stage of B cell growth arrest.*

SLUG and LMO2, anti-apoptotic regulators of normal hematopoietic progenitors^{163,176,177}. This rearrangement has an extremely poor prognosis^{169,178}.

E2A::PBX1 (TCF3::PBX1) arises from balanced translocation t (1;19) (q23; p13) or unbalanced der (19)t(1;19). Both E2A and PBX1 encode transcription factors which when fused forms a protein that aberrantly activates targets such as WNT16 and MerTK¹⁸³⁻¹⁸⁵. The prognosis for pediatric *E2A-PBX1* with current treatment regimens leading to better outcome^{154,178,186}.

ETV6::RUNX1 (t (12;21) (p13;q21) is associated with CD27⁺/CD44^{low-neg}. ETV6::RUNX1⁺ B-ALL has an excellent short-term outcome which cannot be identified based on standard prognostic factors^{187,188}. However, this mutation has been found in relapse¹⁸⁹.

Hypodiploidy (< 46 chromosomes) has a low prevalence in children and consists of three separate subgroups defined based on number of chromosomes which is relevant for disease outcome. Near-haploid (<30 chromosomes) which occurs more frequently in females, with most of the children being younger and low hypodiploid (33-39 chromosomes) are associated with poor prognosis while high hypodiploid (42-45 chromosomes) is the most prevalent of these subgroups^{163,190,191}.

MLL translocations also known as KMT2A rearrangement involve the myeloid/lymphoid or mixed lineage leukemia gene (*MLL*, *TRX*, and *HTRX*) located at chromosome 11q23. MLL fusions are diverse with the possibility of > 70 different translocations, most frequent of which are *AF4*, *AF9*, *ENL*, and *AF6*¹⁶³. MLL-AF4 (common in patients <1 year of age is the most common while MLL-AF9 (t(9;11)(p22;q23)) frequently occurs in secondary malignancies induced by therapy¹⁶³. MLL translocations are high risk, with very bad prognosis (5-year EFS of 30-40%)^{159,163}.

Intrachromosomal amplification of chromosome 21 (iAMP 21) is characterized by chromosome 21 instability and is a high risk subtype¹⁹². The median age of patients is 9 - 11 years, with a poor prognosis that can be improved by chemotherapy intensification^{193,194}.

B-ALL and ALL affects people of all ages - children (0 – 14 years), adolescents and young adults (15 – 24 years) and adults (>25 years) but at different rates¹⁵⁴. Of all age groups, pediatric patients fare better than other age groups¹⁵⁴. Overall, B-ALL in children occurs at a rate of ~25% in Canada and globally, making it the most prevalent pediatric cancer and cause of childhood cancer related deaths¹⁹⁵. Furthermore, pediatric B-ALL is a heterogenous disease, with complexities to continue to pose questions in the field. Thus, pediatric B-ALL is a significant burden.

1.3.3 Pediatric B-ALL treatment

Part of the reason for the difference in the success rate of treatment between pediatric and adult B-ALL is the difference in the treatment regimens. Treatment for HM usually involves one or a combination of multi-agent chemotherapy, radiation therapy, and/or HSC transplant. Newer targeted therapies, such as small molecule inhibitors or monoclonal antibodies, are available to treat B-ALL. Recently, the FDA has approved use of chimeric antigen receptor T (CAR-T) cells that are engineered to target antigens present on tumors. This has allowed for significant improvements of, for example, relapsed or refractory B-ALL, where CAR-T cells that target the pan-B cell marker CD19, can induce remission in 81% of patients after 3 months of treatment¹⁹⁶. The standard ALL treatment based on the Berlin-Frankfurt-Munster (BFM) backbone takes between 2 – ~3 years and occurs in phases –the steroid pre-phase, remission induction, consolidation, and maintenance^{154,182,197,198}. To prevent relapse due to central nervous system (CNS) sequestered lymphoblasts, intrathecal therapy is prophylactically administered during the

consolidation phase¹⁹⁹. Most traditional drugs used for ALL treatment were developed over five decades ago. Thus their dosage and regimens have been stringently optimized which have led to higher survival rates²⁰⁰. For very high-risk patients, allogeneic hematopoietic stem cell transplantation is also considered standard of care²⁰¹.

Steroids are cytotoxic to leukemic cells; thus, they are used for quick clearance of the abnormal B lymphoblasts overcrowding in the BM. Steroid pre-phase begins almost immediately after diagnosis, using steroids such as dexamethasone or prednisone (glucocorticoids - GC). GCs are an essential part of ALL treatment that function by binding cytoplasmatic GC receptors, forming dimers, that can translocate to the nucleus²⁰²⁻²⁰⁵. Within the nucleus, the GC receptor dimers interact with GC response elements to activate gene expression in trans. Alternatively, these receptors function as monomers that repress activity of TF such as AP-1 or NFκB²⁰³⁻²⁰⁵.

The first treatment phase is the remission induction, which uses chemotherapeutic agents such as GC, daunorubicin, vincristine and L-asparaginase, with or without anthracycline^{200,206}. For standard-risk patients, this treatment regimen and intensified treatment post-intensification is usually sufficient; while high and very-high risk patients require a cocktail of drugs including those previously alluded to²⁰⁰. In 96-99% of pediatric and 78-92% of adult patients, remission induction can eliminate the initial lymphoblastic burden, with restoration of normal hematopoiesis as seen by within normal range BM microscopy and WBC levels²⁰⁷⁻²¹⁰. Because chemotherapeutic drugs are non-selectively cytotoxic, they can lead to patients becoming immunocompromised due to death of immune cells. To protect against these cytotoxic effects, drugs to protect against kidney damage, antibiotics and or blood and platelet transfusions may be administered. Due to the possibility of migration of the lymphoblast to the CNS, some patients are treated with intrathecal chemotherapy, with the goal of preventing or treating CNS involvement. Post induction, BM

biopsy and lumbar puncture are done, to assess for MRD. BM⁻ and CNS⁻ status is indicative of remission, with a continuance to the next phase. However, BM⁺ and CNS⁺ status leads to a change in the treatment protocol. In addition to chemotherapy, TKIs are also included in treatment of eligible subtypes specifically Ph⁺ patients¹⁵⁴.

The aim of the consolidation and intensification phases is to eliminate all lymphoblasts and prevent relapse. The specific protocol for this treatment phase depends on the presence or absence of MRD, the type of genetic alterations present and risk stratification. Multi-agents used include methotrexate, chemotherapeutics such as Vincristine, Etoposide, Cytarabine, Leucovorin (to treat methotrexate toxicity) with or without TKIs such as Nilotinib, and anthracyclines such as Daunorubicin^{154,211,212}. In patients who have previously relapsed, especially multiple times, multi-agent toxicity may be a limitation²¹². Thus, there may be attempts at single agent salvage therapy using agents such as vinorelbine, clofarabine, nelarabine and topotecan²¹³. Newer agents include LT3 inhibitor lestaurtinib, BCL2 inhibitor ABT-263, Aurora kinase inhibitor MLN837 and monoclonal antibodies targeting markers such as CD19 (SAR3419, XMAb5574), CD20 (rituximab), CD22 (epratuzumab), CD33 (gemtuzumab) and CD52 (alemtuzumab) and Chimeric antigen receptor (CAR) T cell therapy²¹⁴⁻²¹⁹.

1.3.4 Pediatric B-ALL relapse and Measurable residual disease (MRD)

Relapse is a major cause of cancer related deaths in pediatric B-ALL. Despite the success of therapy after initial diagnosis, 15-25% of patients suffer relapse which reduces survival rates to 30-40%^{220,221}. Relapse can occur in the BM or outside the BM (extramedullary)²²¹. Extramedullary sites including the CNS and the testes are called sanctuary sites, as cells at these sites are partially protected from the effects of chemotherapy^{222,223}. A prognosticator of B-ALL relapse is MRD; low level leukemic cells (<1-5%)^{200,224,225}. Currently, MRD is detected using invasive BM aspiration,

followed by real time quantitative reverse transcriptase PCR (RT-PCR) or FC based analyses^{226,227}. RT-PCR-based analysis is laborious and requires specialized training and tools including patient specific primers and probes^{228,229}. FC based analyses have high sensitivity, depending on the number of fluorophores in the panels²²⁹ and requires specialized expertise for interpretation. Furthermore, for sensitivity of $>10^{-4}$, RT-PCR and FC based analyses require collection of $> 2 \times 10^6$ and $> 5 \times 10^6$ cells²³⁰. Therefore, updated pipelines that reduce the workflow, to facilitate more real time monitoring of MRD and tumor burden may be beneficial. Based on the timeframe from diagnoses, relapse is classified into very early relapse (<18 months after diagnosis), early relapse (> 18 months post diagnosis; but < 6 months after cessation of treatment), and late relapse (> 30 months after diagnosis but > 6 months after treatment cessation)^{154,221}. Early relapses are thought to arise from the selection of a highly dynamic clone of leukemic cells present at diagnosis and persisting despite treatment^{221,231,232}. Typically, these resistant clones are undetectable using conventional methods. In contrast, late relapse may be due to *de novo* development of a second leukemia clone which arose from a common pre-malignant clone^{221,231}.

1.3.5 Other Hematological malignancies

Lymphomas are divided into Hodgkin's (HL) and Non-Hodgkin's lymphoma (NHL), where 90% of lymphomas are NHL²³³. HL is characterized by the presence of multinucleated Reed-Sternberg cells²³⁴. NHL is a heterogeneous group of lymphomas, classified by the site of cell of origin. The majority (85 – 90 %) of cases arise from B lymphocytes with the remaining arising from T or Natural Killer (NK) lymphocytes²³³. In two-thirds of cases, lymphoma presents as a painless lymphadenopathy, usually involving a cervical lymph node, that can undergo rapid expansion in size. Constitutional symptoms, such as fever, anorexia, and weight loss, are also common^{233,235}.

The initiating events for a substantial percentage of HM, particularly those of lymphoid origin, are chromosomal translocations ²³⁶⁻²³⁸. For example, a common cause of Burkitt's lymphoma is a chromosomal translocation between the promoter for the immunoglobulin heavy chain (IgH) and the proto-oncogene c-Myc that results in dysregulated expression of c-Myc and subsequent transformation of cells with an active IgH promoter ²³⁹. Another example is the translocation between chromosome 9 and 22 [t(9;22)(q34.1;q11.2)] generating the Philadelphia Chromosome that encodes the BCR-ABL1 fusion protein, which is a constitutively activated tyrosine kinase and the causative event in 95% of CML cases ²⁴⁰. Mutations of proto-oncogenes and tumor suppressors are also causal for some cases or types of HM ¹⁸⁰. For example, in 80% of cases of chronic myelomonocytic leukemia the genes for SRSF2, TET2, and/or ASXL1 are mutated ²⁴¹. Identification and characterization of genetic alterations have been instrumental in improving diagnosis using molecular criteria as well as for developing targeted therapies, such as the tyrosine kinase inhibitors that are used in CML.

In addition to the molecular identification of genetic rearrangements and mutations, diagnosis of leukemia is based on morphologic, and immunophenotypic criteria of leukemic cells identified in the blood, BM and CSF; further risk stratification also includes determination of response to treatment ²²⁰. Lymphoma is detected initially through physical detection of lymphadenopathy. However, disease confirmation and classification require analysis of LN biopsies at the cellular and molecular levels. Disease staging of lymphoma is based on disease location (within the LN or extra nodal) and extent of disease involvement (i.e. number of LN and/or BM involved) ²⁴².

1.4 Extracellular Vesicles (EVs)

Extracellular vesicles (EVs) are membrane-enclosed nanoparticles released by cells into the extracellular milieu. EVs consist of a lipid bilayer exterior, which is similar to the plasma membrane lipid bilayer with associated cytoplasmic proteins²⁴³⁻²⁴⁵. They are ubiquitous in prokaryotic and eukaryotic cells^{246,247}. Based on their size and mode of biogenesis, they are classified into three broad groups (**Fig 1.7, Table 1.4**). The smallest EVs are exosomes, which range between 30 and 150 nm in diameter²⁴⁷⁻²⁵⁰ and are released via the exocytosis of intraluminal vesicles (ILVs) encapsulated within multivesicular bodies (MVB), synthesized through the endocytic pathway^{251,252}. Exosomes are enriched in tetraspanins (CD9, CD81), endosomal complex required for transport (ESCRT) proteins (Tsg101, Alix), Rab GTPases (Rab27a, Rab27b, Rab11), cholesterol and ceramide^{251,252}. Microvesicles (also called MV, shedding microvesicles, microparticles or ectosomes) are 100 -1,000 nm in size and synthesized via calcium-dependent outward budding of the plasma membrane (PM) in the presence of GTPases RhoA and Rho-associated coiled-coil containing protein kinase (ARF6), resulting in the exposure of phosphatidylserine (PS) on the surface of microvesicles^{253,254}. Also, a tetraspanin - CD63 is present on Microvesicles²⁵⁵. Apoptotic bodies are the largest in the EV subgroups (between 1000 - 5000 nm) and are the product of the final stages of apoptosis²⁵⁶. Less universal type of EVs called oncosomes ((between 1000 - 5000 nm) are exclusively secreted by cancer cells. Exosome and MV-enriched proteins and lipids serve as EV markers. Apoptotic bodies produced during apoptosis are not included in the scope of my studies, thus in this thesis, the term EV refers to exosomes and microvesicles.

Since the first appearance of EVs in the scientific literature in 1987²⁵⁷, EVs have been isolated from diverse cells, organs and tissues using multiple techniques. In the absence of standards of

EVs have been recognized as important constituents of complex network of cellular signaling, which act as vehicles for communication²⁵⁸. EVs transfer their parent cell derived cargo to the recipient cell(s). This transfer can stimulate cellular reprogramming of recipient cell functions needed to effect change or as a response to stimuli ²⁵⁹. For instance, EV-mediated transfer of

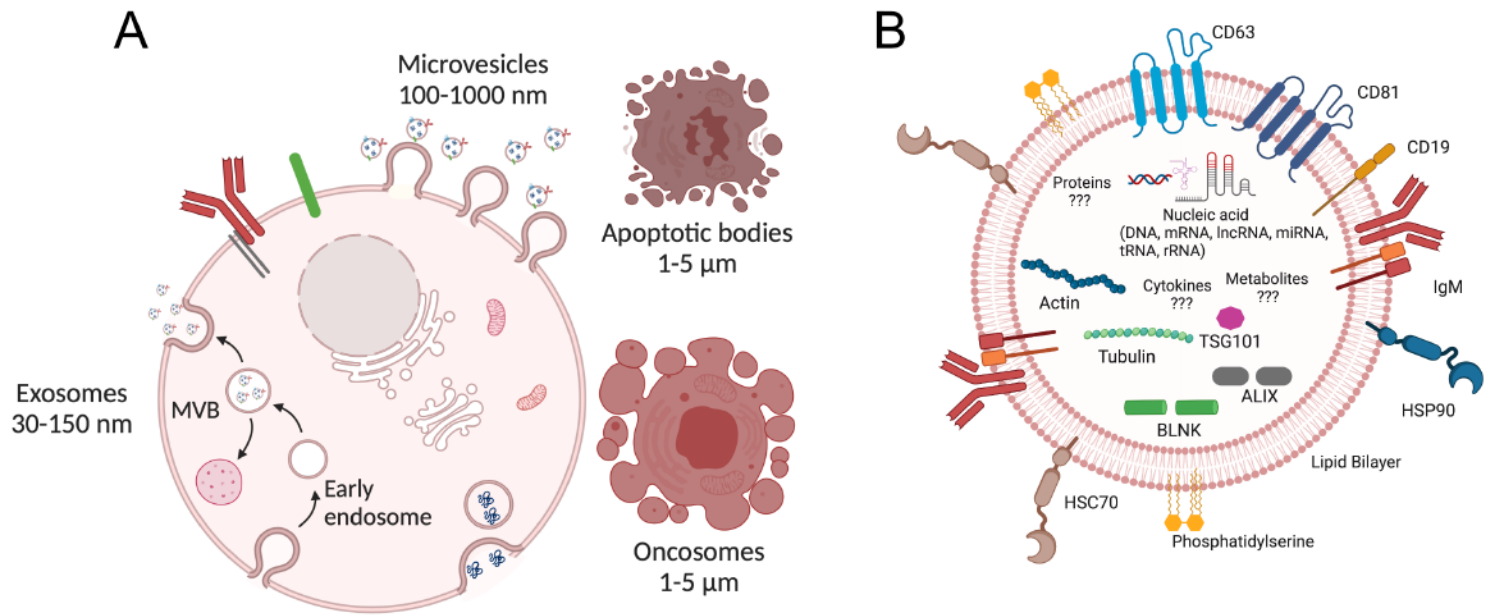


Fig. 1.7. Extracellular vesicles characteristics. A. Extracellular vesicles (EV) are classified into three main types, based on mode of biogenesis and size. B. A sample EV derived from a B cell, showing the bioactive cargo content.

Table 1.4. Extracellular vesicle physicochemical properties and biogenesis pathway.

	Exosomes	Microvesicles	Apoptotic bodies	Oncosomes
Size (nm)	30 – 150	100 – 1000	1000 – 5000	1000 - 10000
Density (g/mL)	1.13 – 1.19	1.032 – 1.068	1.16 – 1.28	unclear
Biogenesis	The endocytic pathway: cell membrane invaginates to form nuclear endosomes (multivesicular bodies – MVB). MVBs buds inwards to encapsulate intraluminal vesicles (ILV). When MVB fuses with the cell membrane, ILVs are released into the extracellular space as EVs.	Cell membrane outward budding and abscission	The shrinking and post-mitotic division of final phase apoptotic cells	The shedding of plasma membrane blebs (non-apoptotic)
Biogenesis mechanism	1. ESCRT dependent pathway 2. ESCRT independent pathway	Calcium dependent pathway	Apoptosis dependent pathway	Cancer cell release, using amoeboid movement

Markers	CD9, CD81, TSG101, ALIX, HSP90	CD63, Integrin, selectin, tissue factor	Annexin A1, histone coagulation factor, C3b	Caveolin-1, ADP ribosylation factor 6
EV cargo	mRNA, miRNA, lncRNA, DNA, protein, lipid, major histocompatibility complex	mRNA, miRNA, DNA, protein, lipid	Proteins, nuclear segments, DNA, RNAs, cell organelles, lipid	Carcinogenic goods, nucleic acids, proteins

information facilitates cell-cell crosstalk between the hematopoietic systems and local or distant cells or tissues ²⁶⁰. Furthermore, EVs are distinct entities whose presence, content, and mechanobiology is linked to the cell of origin ²⁶¹. The release of EVs from cells in physiological and pathological states makes it imperative to study EVs in both conditions. Recent work, as reviewed below, has shown that EVs can distinguish between these two states.

Characterization of extracellular vesicles depends on the successful isolation of EVs, which is made difficult by key features. EVs are within the same size range of other particles including viruses. The presence and/or absence of proteins (surface and/or integral) and lipids (phospholipids and glycolipids) in EVs add another layer of complexity ^{252,262}. These characteristics can be used as basis for the isolation of EVs. Isolation techniques used affect the amount of yield, type and purity of EVs recovered ²⁶³. These techniques can be divided into six broad groups: ultracentrifugation (UCF), density-gradient, precipitation, affinity-based, micro-fluidics selection²⁶⁴ and filtration based method. UCF is a one-step process that uses high centrifugal speed for specific time frame to separate and purify EVs²⁶⁵. The pro of using UCF is that when using centrifugation at different speeds such as low speeds, it is possible to isolate a subset of EVs such as larger EVs from low speeds²⁶⁶. The limitation of UCF is that it is a time-consuming method that requires specialized equipment and results in low yield²⁶⁷. Density-gradient separation is a two-step process that uses a density gradient medium (e.g iodixanol and sucrose) and centrifugal force to isolate EVs based on the buoyant density of EVs^{243,266}. This process leads to high purity but coprecipitates high density lipoproteins with the EVs²⁶⁸. This technique also has a long turnaround time and requires specialized equipments. Repeated centrifugation can reduce the amount of non-EV particles co-isolated with the EVs, however, it also results in reduced particle yield due to lost and damaged EVs. Precipitation based methods is the second most popular method and uses

polymers such as polyethylglycol (PEG) to lead to high EV yield (up to 90%)^{243,269}. This method is inexpensive but can also co-isolate lipoprotein contaminants²⁴³. The affinity-based methods use immunoaffinity capture and affinity and ion-exchange chromatography. Immunoaffinity capture is based on presence of surface proteins on the EVs. Affinity and ion exchange chromatography is based on interaction with immobilized ligands on the stationary phase²⁴³. While this method has high specificity, it has lower yield and is time consuming. The microfluidics-based technique is a low-cost tool that can enable accurate and sensitive sorting of EV subpopulations, which is an emerging and promising technique²⁶⁸. Filtration based methods includes ultrafiltration and size exclusion chromatography, both of which involve multistep processes that lead to good yield. Each technique can be used alone or in combination with another technique for optimum results, as well as to compensate for limitations in each. An example of a precipitation-based isolation is peptide affinity-based isolation, using a synthetic peptide called Vn96. Vn96 is a 27-mer synthetic peptide with an affinity for EV surface-associated heat shock proteins. When incubated with tissue culture media and biofluids, the resultant solution can be centrifuged at low speed to precipitate EV-peptide complexes²⁷⁰. Limitations such as irreversible interaction with the isolation material, potential presence of protein aggregates, and low purity can all influence downstream analysis. There is currently no benchmark for EV isolation, although UCF is considered a “gold standard” for purity but not yield^{271,272}. A major challenge with EV studies is the lack of standard methods to obtain highly pure and functional EVs for downstream analysis²⁶³.

EV characterization and quantification methods include western blotting, single-particle tracking (SPT), electron microscopy, Nanoparticle tracking analysis (NTA), tunable resistive pulse sensing (TRPS), dynamic light scattering (DLS), Atomic Form Microscopy (AFM), and Flow Cytometry^{273,274}. Western blots and flow cytometry are used to profile protein markers using gel-

electrophoresis or fluorophore-conjugated antibodies. The NTA is used to quantify EV size and concentration, while electron microscopy is used to assay EV size and shape. SPT and AFM are used to visualize EV interactions while DLS is used to quantify EV size. Each characterization method analyzes a different property of EVs, which determines the type of data output (linked to downstream applications). Downstream applications include in-vitro functional assays, transcriptomics, proteomics and lipidomics, which explore current knowledge.

1.4.1 EV secretion and cargo

During EV biogenesis, a portion of the originating cell's PM and cytosol is transferred to the vesicle, resulting in bioactive cargo (proteins, lipids, and nucleic acids)²⁵⁹ (**Table 1.4**). Specific mechanisms of EV cargo packaging depend on cargo type, abundance in the cell, and cell physiological/pathological state²⁷⁵. Furthermore, EV cargo dictates EV biology, including EV uptake. EV uptake could occur via direct fusion of the EV to the PM, non-specific endocytosis (pinocytosis), or specific endocytosis, including receptor-mediated endocytosis where EV surface proteins act as ligands to recipient cell surface receptors and facilitate uptake and cargo transfer²⁷⁶. Hence, EV bioactive protein cargo is vital for uptake by recipient cells and downstream effects in cell physiology and pathology-dependent manner. Packaging of selective cargo determines the final EV composition, which in turn dictates the fate and function of EVs²⁷⁷. Hence, EV cargo content can serve as EV biomarkers for identifying specific EV subtypes and cell of origin²⁷⁸⁻²⁸¹. EVs can transfer bioactive cargo from the donor to recipient cells to mediate downstream effects. Proteins account for a substantial amount of EV bioactive cargo. EV protein cargo is in part dependent on the cell of origin and the mode of EV biogenesis²⁸². Studies by different groups suggest that protein packaging into EVs occur via the interaction of the proteins with components

of the EV biogenesis machinery^{283,284}. For instance, exosomes (synthesized via the endocytic pathway) tend to be enriched in proteins such as MHC class II, tetraspanins (CD37, CD53, CD63 and CD82)²⁸⁵; endosomal sorting complex required for transport (ESCRT) and the associated proteins (Alix, TSG101)²⁸⁶ and chaperones such as Hsp90^{287,288}. On the other hand, microvesicles which are budded off the cell's plasma membrane are enriched with different sets of plasma membrane linked proteins. Microvesicles often contain proteins such as integrins, glycoprotein Lb (GPIb) and P-selectin²⁸⁵. Hence, protein cargo identification can be used for EV typing.

EV proteins could be attached to the surface (surface proteins), or embedded in the EV core²⁸⁹ depending on the mode of biogenesis and release of the EVs. In addition, EV surface proteins may provide information about the biological state of the parent/secreting cell²⁸². For instance, antigenic profile of the expression of proteins such as CD24, Ca-125, CA19-9, EGFR and claudin 3 in cancer cells and the protein cargo profile of released EVs were consistent²⁹⁰. Hence, EV protein cargo content can be used for predicting parent cell protein content.

EVs are bound by a phospholipid bilayer, with lipid EV content including plasma membrane lipids. Lipids in EVs exist either alone or in conjugation with other biomolecules such as in glycolipids, phospholipids, and cholesterol. EVs are enriched in lipid rafts (microdomains) which are membrane subdomains enriched in cholesterol, saturated fatty acids and sphingomyelin with a tightly packed and ordered density²⁹¹ which are important for EVs biogenesis and release.

EVs contain a unique array of nucleic acids as part of their cargo^{292,293}. The RNA content of extracellular vesicles differ for EV types and parent cells²⁹⁴. Current evidence suggests that miRNA shuttling into EVs use specific mechanisms, which may be cell-specific, with a general model of miRNAs forming complexes called miRNP with RNA binding proteins (RBPs) such as SYNCRIP, Alyref, Fus^{295,296}. Specific signals on miRNAs such as specific motifs in their sequence

target them for cell retention versus export into EVs^{297,298}. Different RNA species including mRNA and small ncRNA are selectively packaged into EVs during their biogenesis²⁹⁹. EVs' nucleic acid cargo can be transferred between cells. EVs can also transfer genomic DNA (gDNA) and mitochondrial DNA (mtDNA) from parent to recipient cells³⁰⁰; in a manner that increases the gDNA coding mRNA and protein levels and inducing new functions in recipient cells³⁰¹.

The underlying mechanism behind EV secretion is one of the open questions in the EV field. For instance, with exosome release via the endocytic pathway, it is not known what stimuli sends the cell down the path that culminates in MVB fusion with the PM as opposed to the lysosomal degradation pathway. However, some studies have identified that stress may be a trigger for EV secretion. Different forms of cellular stress such as endoplasmic reticulum (ER) stress, hypoxia, nutrient deprivation, hypoxia, heat stress and oxidative have been linked to EV release. Specifically, ER stress was shown to be associated with formation of MVB and increased EV release in HeLa cells³⁰². In a broad range of cancer cell types including leukemias, different types of cellular stress was shown to be associated with increased EV secretion from these cells³⁰³. Thus, we can surmise that cell stress modulates EV release.

Another downstream effect of cell stress is the induction of autophagy - a process that mediates metabolic adaptation to stress, to protect against potential damage. Autophagy is a highly conserved and selective clearance pathway that leads to the degradation of defective organelles and misfolded proteins using lysosomes, in order to maintain cellular and tissue homeostasis³⁰⁴. In autophagy, cells form double-membraned vesicles called autophagosomes, which hold sequestered organelles, proteins, or cytoplasmic portions, that are delivered to the lysosome for degradation. Autophagy occurs in five main stages: initiation, nucleation, elongation and maturation, fusion and degradation³⁰⁵. Each stage is regulated by different signals and associated

pathways. Initiation is regulated by mTOR (positive regulation by Akt and MAPK signals inhibit autophagy while negative regulation via AMPK and P53 signals activate autophagy)³⁰⁶. Autophagosome formation is controlled by autophagy-related (ATG) proteins including the ULK complex, phosphatidylinositol 3-kinase (PI3K) complex, the Atg2-Atg18/WIP1 complex, Atg9 and the conjugation systems (Atg12 and Atg8/LC3)³⁰⁷. During autophagosome formation, Beclin-1 (which regulates autophagosome) complexes with Vps34 and Atg14L to promote nucleation^{305,308}. The elongation and maturation stage is regulated by LC3 and Atg12-Atg5-Atg16 complex³⁰⁵. Together, autophagy is regulated by a complex set of molecules and proteins.

A link between EV biogenesis and autophagy exists, wherein EVs from some cells can activate or inhibit autophagy by targeting autophagy associated molecules and pathways. For instance, EVs from non-small cell lung carcinoma cell line A549 are enriched in miR-425-3p which target Akt/mTOR pathway, to activate autophagy³⁰⁹. Human bronchial epithelial cells derived EVs carry miR-7-5p, which inhibited the Akt/mTOR pathway leading to autophagy³¹⁰. Human neuroblastoma cell derived EVs are enriched in miR-19a-3p which regulated the Akt/mTOR signalling pathway this inhibiting autophagy³¹¹. Furthermore, exosomes carrying miR-30a (regulator of Beclin-1) have been shown to be packaged into EVs, which inhibited autophagy, to reduce cell apoptosis in cardiomyocytes^{312,313}. Thus, by carrying miRNAs that regulate expression of autophagosome pathway proteins and molecules, EVs can modulate cell autophagy.

There is a crosstalk between autophagy and exosome biogenesis at different stages of autophagy and exosome biogenesis, respectively. Autophagy and exosome biogenesis pathways intersect at different stages, and share certain proteins and markers³¹⁴. Specifically, EV and autophagy flux are both regulated by common molecules. SNARE family proteins such as VAMP7, and syntaxin 7 and 8 facilitate MVB fusion to PM and mediates the fusion stage of autophagy^{315,316}.

Additionally, during the autophagosome assembly phase, an MVB protein Rab11 acts as a platform for ATG proteins³¹⁷. EV cargo protein ALIX has been suggested to be linked to autophagy, with ALIX inhibition decreasing autophagy²⁴⁹. Thus, SNARE proteins and other exosome biogenesis proteins are a link between autophagy and exosome biogenesis. Conversely, autophagy related proteins are linked to regulation of EV biogenesis at normal and physiological states. Different ATG proteins such as ATG5, ATG9 and ATG12-ATG3 have been implicated in regulating ILV and MVB synthesis at different stages of the endosomal pathway, as well as being included as EV cargo³¹⁴. ATG5 and ATG16L1 deficient breast cancer cells show reduction in EV release while G alpha interacting protein C-terminus (GIPC) induced EV release^{318,319}. Other regulators of the autophagy phases, such as ATG12-ATG3 complex can contribute to exosome biogenesis by interacting with the ESCRT machinery protein ALIX that produces ILVs³²⁰. In CML cells, the distribution of PI3K which facilitates membrane trafficking in endocytosis and autophagy, led to reduction of exosome biogenesis and autophagy^{321,322}. Furthermore, in some cell contexts autophagy and exosome biogenesis compensate for each other^{314,323}. In mouse study models, blocking exosome secretion via ISGylation of TSG101 directs MVBs to the autophagy pathway, while inhibition of autophagy (by blocking lysosome-endosome fusion) restores exosome secretion³²⁴. Furthermore, MVBs that cannot be degraded can be diverted to the autophagy pathway, while aberrant autophagy can direct MVBs to the PM and stimulate release of exosomes³²⁵. Specifically, during exosome biogenesis, mTORC1, a regulator of autophagy regulates exosome release, in a Rab27A dependent manner³⁰⁶. Autophagy and exosome biogenesis (at the MVB stage) are concurrently regulated by mTORC1, which allows cells to coordinate waste management and recycling³²⁶. Thus, it is possible that at physiological conditions, autophagy uses exosome biogenesis as an alternate route for biomolecule recycling and shuttling to proximal and

distal cells. Conversely in pathological conditions such as cancer, insufficient autophagy occurs, which may be linked to increase in EV release as seen in many cancers³⁰⁶. In turn, cancer cell derived EVs can induce autophagy in recipient cells, which stimulates growth, migration and enhances drug resistance³¹⁴. Furthermore, it is well documented that in cancer cells, strong activation of both the autophagy and EV biogenesis pathways are part of the cancer cell response³¹⁴. Taken together, stress induced autophagy and EV biogenesis drive each other in a way that regulates cargo packaging and in turn facilitates normal physiology or cancer survival and progression. Therefore, the exosome and autophagy pathways are synchronized pathways for intracellular removal, with each pathway able to compensate for each other.

1.4.2 EVs and Liquid biopsy

In traditional disease monitoring, tissue biopsy is the standard for diagnostics and monitoring therapy response. Tumor molecular profiling depends on tumour sampling, possible through invasive procedures including collection of resected tumor samples and tumour biopsies. The main limitations of these current methods includes difficulty in obtaining adequate tumor quantity and quality; and the invasiveness of traditional biopsy making it difficult to sample real-time especially as frequently as is needed to monitor therapy response and during relapse^{327,328}. Furthermore, tumor heterogeneity and tumor drift (both spatially and temporally) in response to treatment creates additional layers of complexity with using traditional tumor biopsies^{327,329}. Tissue biopsy-based diagnostics have limitations in disease assessment such as inability to capture tumor heterogeneity, the invasiveness of the sampling makes it impractical for monitoring tumour response and relapse^{327,330}. However, these limitations have led to unsatisfactory monitoring, which necessitates the search for more novel, revolutionary approaches.

In recent years, cancer patient outcomes have improved tremendously due to tumor molecular profiling. To circumvent these limitations, liquid biopsy is emerging as a less invasive tumor sampling method, hinging on the premise that biofluids such as blood, mucosa, pleural effusions, urine and cerebrospinal fluid contacts most tumors³³¹. Thus, biofluids like blood carry tumor artefacts, some of which can be used as biomarkers for disease monitoring.

Liquid biopsy typically involves an alternate sampling method but does not completely replace the traditional approach(es). Liquid biopsy analytes (**Fig. 1.8**) varies between diseases, but in cancer include biomarkers such as circulating tumour DNA (ctDNA), cell free DNA (cfDNA), disseminated tumor cells (DTCs), cfRNA, tumor-derived factors, circulating protein biomarkers and EVs³³². EVs have been shown to be isolated intact from biofluids, for liquid biopsy, with EVs used in lieu of cells as carriers of actionable biological information^{333,334}. This makes EVs a good alternate to solid tumor biopsy. An advantage of EVs over other liquid biopsy analytes is that EV content are protected by the lipid bilayer membrane³³⁵. Thus, EV associated cargo whether still in circulation or after being isolated which leads to EVs being more stable sources of biomarkers. In addition, EVs are stable across different cold storage conditions, with -80°C being the most stable for long periods³³⁶. Thus, EVs can allow for liquid biopsy from remote areas, since biofluids or even isolated EVs retain their bioactive characteristics despite cold chain transportation over long distances and time periods including biobanked samples. Another strength of EVs as liquid biopsy biomarkers is their ubiquity in body fluids³³⁷⁻³⁴⁰. Their ubiquity (especially in normal versus disease systems) allows for clear comparisons, to identify signatures that could drive quicker monitoring. The ubiquity and heterogeneity of EV cargo will also allow for potential multiparametric analysis, which can provide an accurate longitudinal snapshot of the disease at different timepoints^{341,342}. Furthermore, in human malignancies, tumors exhibit heterogeneity at

two levels. First, there is inter-tumour heterogeneity – tumors of the same histopathological subtypes differing from one patient to another; and intra-tumour heterogeneity - variations within

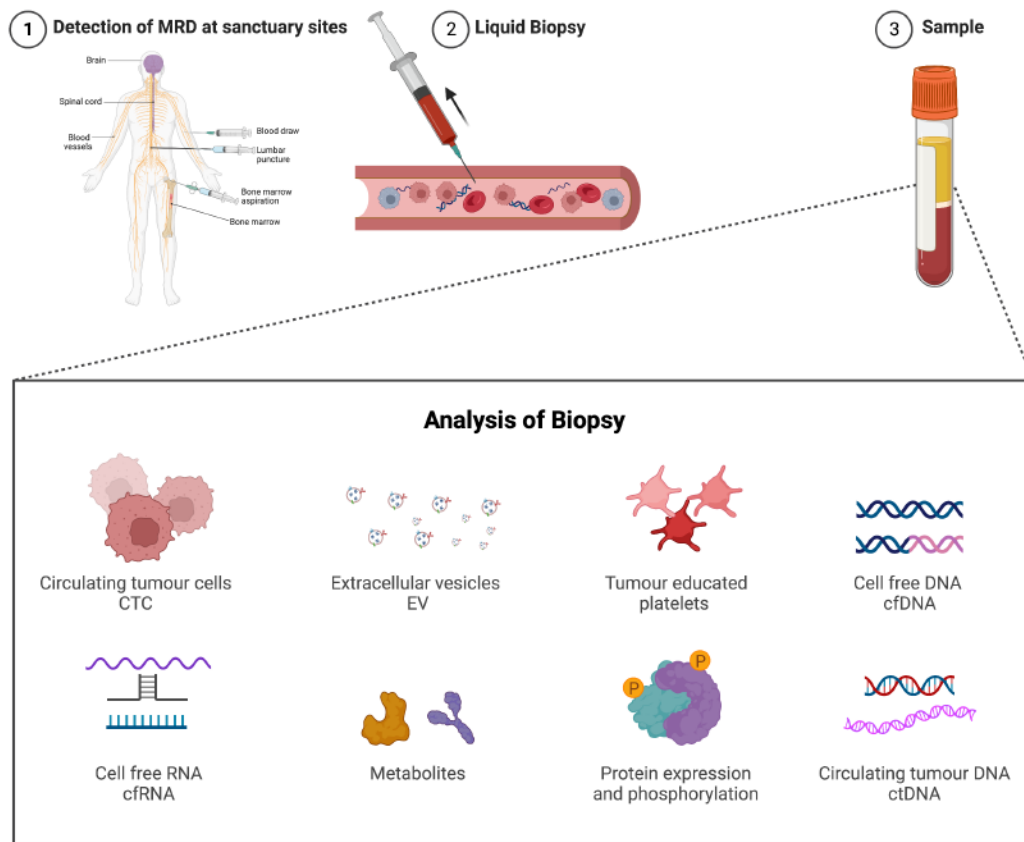


Fig. 1.8. Comparison of tumour biopsy techniques. Liquid biopsy approaches including extracellular vesicles, circulating tumour DNA (ctDNA), and cell free DNA (cfDNA).

a single tumour³⁴³. Tissue biopsy is unable to capture any of these heterogeneities, as it is impractical to sample enough portions of the tumour. Aspects of EVs that can be used as liquid biopsy biomarkers include EV size, concentration in biofluids, and cargo. EVs can be used for real-time monitoring of therapy response and resistance in patients, because of the dynamic nature of circulating EV and their cargo^{344–347}. Overall, EVs are emerging as the next big thing for liquid biopsy – a less invasive alternative to tissue biopsies that can allow multiparametric analysis, giving a holistic picture of the disease landscape, at both single, longitudinal, and cross-sectional timepoints.

1.4.3 Knowns and unknowns about extracellular vesicles

The current gold standard for EV isolations – UCF is not very feasible for EV isolations in clinical context. The currently emerging multiplexed approach (e.g SEC and dUCF) though compatible with clinical contexts raise other issues including isolating only EV subpopulation (based on markers). Key criteria for considerations include sample nature, volume, purity and downstream analysis³⁴⁸. Though these considerations are known to all, there are a few guidelines which are not robust, – mostly because the EV field is still relatively nascent. Beyond considerations for EV isolation, the methods of analysis of EV cargo are also currently unstandardized.

1.4.4 EVs in the regulation of normal immune function

Normal immune cells use EVs for cell-cell communication for immune response and to maintain homeostasis. All immune cells tested secrete EVs of diverse sizes and cargo content. To appreciate the potential of EVs in HM, it is important to understand their role in the immune system.

By transferring immune regulatory proteins to recipient cells, EVs influence, and mediate activation and progression of events involved in the immune response. For instance, under

inflammatory conditions, THP-1 macrophages synthesize and release microvesicles at high concentrations, which can induce activation, phenotypic differentiation, and maturation of naïve monocytes into macrophages via the transfer of *miR-223*³⁴⁹. Monocytes can cause inflammation-mediated apoptosis in smooth muscle cells by transferring caspase-1 via microvesicles³⁵⁰. In addition, EVs package endogenous pro-inflammatory ligands, such as high-mobility group box proteins (HMGB) -1 or -2^{351–353}, that can activate innate and adaptive cells of the immune system. Packaging of pathogen-associated molecular patterns (PAMPs) (e.g. glycopeptidolipids) and/or pro-inflammatory cytokines (e.g. IL-1) can promote inflammation in recipient cells^{354,355}. Likewise, packaging of anti-inflammatory cytokines (e.g. TGF- β) can restrict these responses in target cells³⁵⁶. Packaging of pro- or anti-inflammatory cytokines is dependent on the status and identity of the originating cell. In the context of the adaptive immune response, EVs can be taken up by Antigen Presenting Cells (APC), which can load the protein cargo from the EVs onto MHC II for presentation to CD4⁺ T helper cells. APCs can also release EVs that carry MHC/antigen complexes and associated co-stimulatory molecules to directly activate CD4⁺ T cells^{357,358}. Thus, during either the innate or the adaptive immune response, EVs act as paracrine messengers that can induce or propagate inflammatory signals or negatively regulate inflammation. Steady-state (basal) secretion of EVs by HM cells can impair the normal immune response. For example, EVs isolated from the T-ALL cell line CCRF-CEM altered expression of inflammatory genes in human monocytes, an effect modulated by the presence of tumor necrosis factor (TNF- α)³⁵⁹. EVs released by Waldenstroms' macroglobulinemia (WM) cells can package mutant MyD88, a driver protein in WM biogenesis, which can be taken up by mast cells and macrophages resulting in activation of proinflammatory signaling³⁶⁰. Exosomes released from Epstein Barr Virus-infected Burkitt's lymphoma cells can induce proliferation of normal primary B cells and promote B cell

differentiation³⁶¹. CLL-derived exosomes stimulate the adoption of an immunosuppressive phenotype in monocytes due to the upregulation of the checkpoint inhibitor PD-L1³⁶². This process is dependent on activation of Toll-like receptor 7 (TLR7) by the small ncRNA hY4 packaged in CLL EVs. Thus, the release of EVs by HM into the microenvironment or systemic circulation can directly impact the proper functioning of the immune system. Receptor-induced release of EVs can modify the microenvironment in a manner that is dependent on the extracellular milieu. Stimulation of the BCR which normally detects the presence of foreign antigen, has been shown to increase the number of exosomes secreted by B lymphoma cells^{363,364}. Moreover, this stimulation alters exosome cargo with increased levels of BCR and MHC class II as well as increased packaging of *miR-150* and *miR-155*³⁶³⁻³⁶⁵. In a similar manner, stimulation of the CD24 receptor in a lymphoma cell line increases the release of total microvesicles and CD24-bearing microvesicles³⁶⁶. In addition to packaging the BCR and CD24 on EVs, stimulation of these receptors induce packaging of other surface receptors on the EVs^{351,364}. Interestingly, stimulation of CD24 changes the abundance of some of these receptors on microvesicles³⁵¹. The implications of changes in surface protein composition are not yet known but we speculate that they may affect EV recipient cells.

1.4.5 EVs in immune evasion

Immune surveillance for cancer cells is a complex process orchestrated by numerous types of immune cells; however, natural killer (NK) cells have a primary role in this process that can be affected by EVs. NK cells are controlled by a balance of activating and inhibitory receptors that differentiate healthy cells from cells stressed by insults such as oxidative stress or genotoxic stress³⁶⁷. EVs secreted by cancer cells can interfere with NK-cell based surveillance, allowing cancer cells to evade detection and destruction. NKp30 is an NK cell activating receptor from the

natural cytotoxicity receptor (NCR) family³⁶⁸. BAG6 is an activating ligand for NKp30 that is released on EVs, in a p53-dependent manner, in response to cellular stress^{369,370}. Exosomal BAG6, but not freely soluble BAG6, was found to increase NK cytolytic activity towards solid tumors³⁷⁰. Interestingly, a reduction in BAG6-expressing exosomes was observed in plasma from CLL patients, which was accompanied by a concomitant increase in soluble BAG6, thus resulting in decreased NK-mediated killing of CLL cells³⁷⁰.

In contrast, Natural Killer Group 2, member D (NKG2D) is an activating NK cell receptor and, after engagement by its ligands, forms a cytotoxic system to eliminate transformed cells³⁶⁷. Thermal and oxidative stress of T-cell leukemia and Burkitt's lymphoma cell lines can enhance the secretion of exosomes bearing NKG2D-ligands³⁷¹. These NKG2D-ligand bearing exosomes decrease NKG2D-dependent NK cell cytotoxicity, contributing to immune evasion of cancer cells, potentially by acting as decoys for the NK cells³⁷¹. AML-derived exosomes contain higher levels of transforming growth factor beta-1 (TGF- β 1), which are then able to inhibit NKG2D receptor expression in NK cells in a manner that correlates with TGF- β levels³⁷².

The presence of EVs that express NK ligands can have major clinical consequences. For example, a recent phase 1 clinical trial looking for immunological recovery and response from the injection of NK-92 cells (human IL2-dependent NK cells) into relapsed AML patients was unsuccessful due to interference by AML exosomes³⁷³. The AML EVs from these patients expressed multiple inhibitory ligands. When NK cells interacted with these exosomes, there was less expression of pro-inflammatory chemokines such as CCL3 (MIP-1a), CCL4 (MIP-1b) and CCL5 (RANTES), as well as upregulation of the adenosine pathway, all of which promote disease progression³⁷³. Thus, EVs produced by HM can act as decoys to inhibit targeting of HM by immune cells or

actively alter NK cell function. Whatever the mechanism, the presence of HM-derived EVs can significantly alter the response to therapies that depend on the immune system.

1.4.6 EVs and the bone marrow microenvironment

The BM microenvironment consists of bone marrow stromal cells (BMSCs) that support the growth of HSCs that develop into mature immune cells. EVs released by HM cells can be internalized by BMSCs (**Fig. 1.9**), which in turn alters BMSC function to promote growth of HM via multiple mechanisms. Exosomes from primary AML cells and AML cell lines transfer multiple transcripts to the murine BMSC line OP9³⁷⁴. These include *IGF-IR*, *CXCR4*, *NPM1*, *MMP9*, and *FLT3* mRNAs. Transfer of the transcript for insulin-like growth factor receptor (*IGF-IR*), and *miR-150*, which targets the CXC chemokine receptor type 4 (*CXCR4*), promotes OP9 recipient cell proliferation and migration³⁷⁴. In addition, AML-relevant mutations in these genes were found in the stromal cells after exosome uptake³⁷⁴. Similarly, primary BMSCs treated with AML-derived exosomes had increased expression of genes that support AML growth, such as *DKK1*, *IL-6*, and *CCL3*, and decreased expression of genes that support normal hematopoiesis, such as *IL-7*, *CXCL12*, *KITL*, and *IGF1*³⁷⁵. Furthermore, exosomes from AML inhibit normal hematopoiesis of HSCs by suppression of c-Myb via exosomal miR-155^{376,377}. AML exosomes, via miR-4532, can suppress expression of the leucine zipper downregulated in cancer-1 (*LDOC1*) transcript, which in turn activates the STAT3 signaling and thus *DKK1* expression to inhibit hematopoiesis³⁷⁸. Related to this, treatment of primary AML cells with a *DKK1* inhibitor can delay AML progression and enhance survival in AML-engrafted mice³⁷⁵, thus identifying alterations in EV cargo may suggest potential therapeutic strategies.

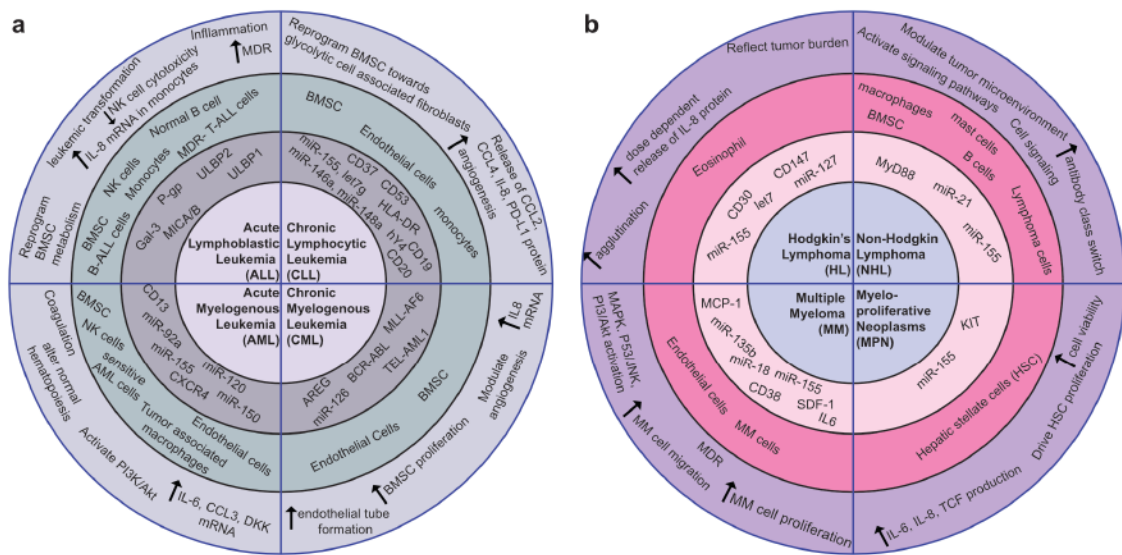


Fig. 1.9. A summary of the bioactive cargo of EVs from hematological malignancies and their downstream effects upon EV uptake by recipient cells. Here, we provide a visual summary of the current state of knowledge for the bioactive cargo and effect of EVs in (a) leukemia and (b) other hematological malignancies (HM). To provide a clearer image of the global impact of EVs from these cells, we have omitted the specific connections between each component. These connections, as well as abbreviations and citations, can be found in Supplementary Table S1. The innermost circle shows the HM that secrete EVs. The next circle shows the cargo, when known, that is present in the EVs secreted by the HM in the matching quadrant. The next circle shows the target cells known to be affected by HM-secreted EVs. The outermost circle shows the effects of the EVs on recipient cells. The increased density in some quadrants reflects the amount of published literature on some HM and not biological differences between HM.

Exosomes derived from primary CML cells and cell lines modulate BMSCs. CML-derived EVs can express AREG, a ligand for epidermal growth factor receptor (EGFR). Activation of EGFR then promotes the secretion of IL-8 by BMSC, which in turn promotes CML proliferation and survival^{379,380}. Also, adhesion of CML cells to BMSCs is enhanced by CML-derived exosomes, which can further support CML survival, likely because of exposure to increased local concentrations of IL-8³⁷⁹. CML-derived microvesicles can also package CML-specific oncogenic fusion transcripts, including BCR-ABL1, and TEL-AML1³⁸¹. BMSCs can then internalize these EVs, which increases BMSC proliferation.

Uptake of EVs from ALL cell lines (SD1 and NALM6) by BMSC significantly reduces mitochondrial respiration and the ability to mount an oxidative response to metabolic challenge and stress in the BMSCs, suggesting that EVs can metabolically reprogram stromal cells toward aerobic glycolysis³⁸². It is likely that this alteration promotes leukemic survival via release of metabolic intermediates from the stromal cells. Furthermore, exosomes can promote chemoresistance by transfer of the carbohydrate binding protein galectin-3 (Gal-3) from BMSCs to ALL in response to cytotoxic drug treatment³⁸³. The increase in Gal-3 in ALL cells, by exosome-mediated transfer and increases in endogenous Gal-3, promotes chemoresistance only when ALL cells are cultured in the presence of stromal cells³⁸³.

Similarly, EVs isolated from the plasma of CLL patients can promote disease progression. For example, CLL-derived microvesicles can maintain activation of AKT-induced production of vascular endothelial growth factor (VEGF) by BMSC, a survival factor for CLL B cells³⁸⁴. In addition, CLL-derived microvesicles increase the expression of cyclin D1 and c-MYC in BMSCs to promote BMSC survival and proliferation³⁸⁴. CLL-derived exosomes can also induce BMSC

to acquire a cancer-associated fibroblast (CAF) phenotype resulting in the activation of multiple pro-proliferative signaling pathways in the BMSC and altering the tumor microenvironment ³⁸⁵.

Adult T-cell leukemia/lymphoma (ALT)-derived exosomes also support the creation of a microenvironment that supports leukemia progression. ALT-EVs contain the leukemia related Tax mRNA and protein, VEGF protein, and miRNAs *miR-21* and *miR-155* ³⁸⁶. Transfer of these cargo to BMSCs induces changes in cell morphology, increases proliferation, and increases expression of genes related to migration and angiogenesis ³⁸⁶.

Exosomes from MM contribute to myeloma bone disease, which is the most common complication in MM and results in osteolytic lesions and severe pain, among other symptoms ³⁸⁷. Myeloma bone disease is due to increased osteoclast activity, which causes bone-resorption, and suppressed osteoblast function, which is responsible for bone formation. Exosomes from MM cell lines and patient serum were shown to activate osteoclast activity by recruiting osteoclast precursors and promoting their differentiation in mature osteoclasts *in vitro* ³⁸⁸. Similar to CML, this effect was found to be due to the activation of EGFR by exosomal-associated AREG, which in turn induced secretion of IL-8 from BMSC to support osteoclastogenesis ³⁸⁹. Exosomes from MM also suppress osteoblast differentiation from BMSC, further impairing bone regeneration, *in vitro* and in a mouse model ^{389,390}. This effect is mediated, at least in part, by the induction of miR-103-3p expression in BMSC ³⁹⁰, however, the targets of miR-103-3p involved in this process have not yet been described. EVs secreted from BMSC in response to the presence of HM or in response to chemotherapeutic agents can also modulate cancer cell survival and progression. As recently reviewed ³⁹¹, BMSC from MM patients (MM-BMSC), but not healthy individuals', promote MM proliferation, survival, and drug resistance *in vitro* and *in vivo*, which is, in part, facilitated by exosomes ^{392,393}. The tumor suppressive miRNA *miR-15a* is lower in exosomes from MM-BMSCs

resulting in increased proliferation of MM cells ³⁹³. Additionally, IL-6 present in MM-BMSCs exosomes was key in promoting tumor proliferation ³⁹³. In addition to transferring cargo, EVs from MM-BMSC activated pro-survival signaling pathways in MM cell lines, including the upregulation of anti-apoptotic BCL-2 and the p38, p53, JNK and Akt signaling pathways ³⁹². EVs derived from BMSC from patients with AML or Myelodysplastic syndromes (MDS) had a significant increase in *miR-7977*, which was subsequently linked to inhibition of normal hematopoiesis via inactivation of the Hippo/Yap pathway in the BMSC ^{394,395}. Moreover, exosomes derived from BMSCs from AML patients were enriched for cytokines, including TGF- β 1, and miRNAs such as *miR-155* and *miR-375*; factors previously associated with increased AML risk ³⁹⁶. Thus, HM EVs can facilitate and promote communication needed for proliferation and survival of HM (**Fig. 1.10**). Furthermore, by altering EVs released by BMSC, the metastatic niche and proliferation signals can be modulated directly, or indirectly.

1.4.7 EVs in angiogenesis, motility, and adhesion

EVs promote angiogenesis, adhesion, and migration. Angiogenesis during the formation of non-circulating tumors increases vascularization around tumor cells to maintain oxygen and nutrient availability. Also, changes to cell motility and adhesion alter the ability of tumor cells to migrate and adhere to secondary sites and can promote the spread of HM cells to naïve sites.

Exosomes derived from leukemias, or lymphomas can transfer pro-angiogenic factors to endothelial cells. For instance, *miR-92a*, packaged in EVs from the CML cell line K562, stimulates migration and vascular tube formation of human vascular endothelial cells (HUVECs) ³⁹⁷. Exosomes derived from K562 cells exposed to hypoxia contained higher levels of *miR-210*, which

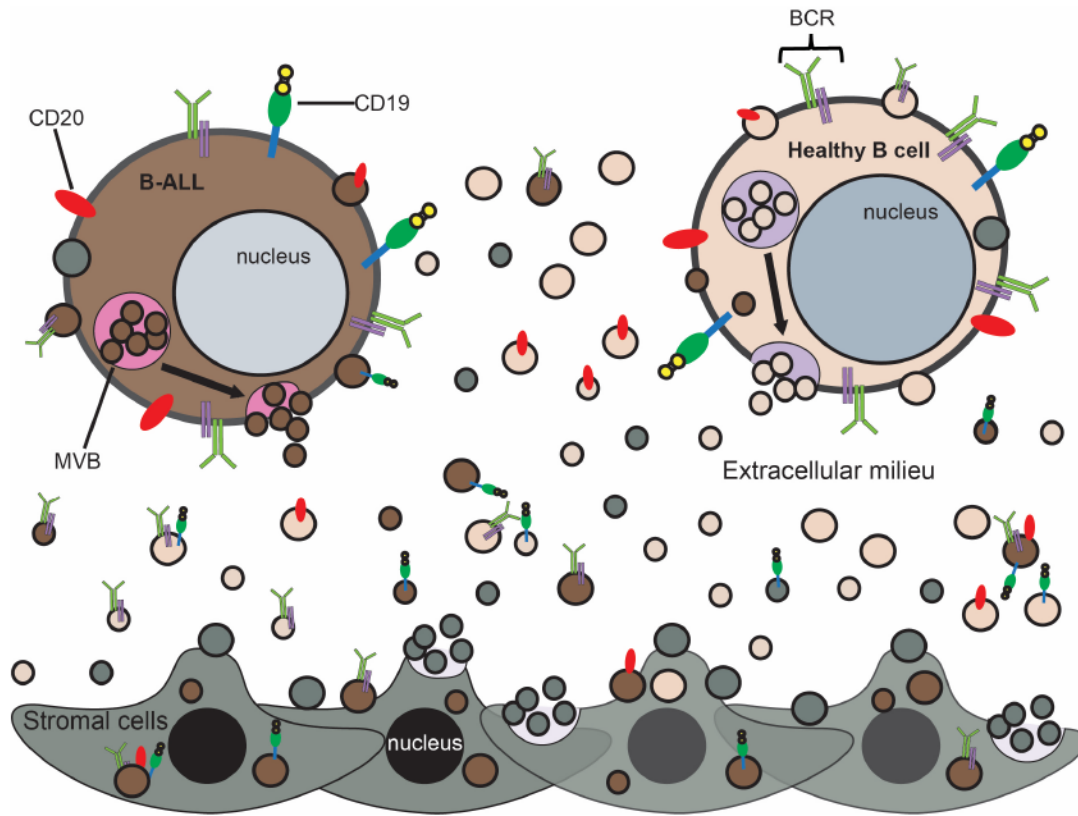


Fig. 1.10. Schematic diagram showing EV-based communication between B cell lymphocytes, B-ALL lymphoblasts, and stromal cells. B cell lymphocytes (beige), B-ALL lymphoblasts (brown), and stromal cells (grey) release exosomes via multivesicular bodies (MVBs) and membrane-budded microvesicles (MVs) into the extracellular milieu. Donor cells and their respective EVs are shown in the same color. All three cell types can take up and incorporate EVs and EV cargo as represented by the differently colored “patches” in each cell. Also depicted is the generation of exosomes and MVs. Exosomes originate as intraluminal vesicles within pink/purple/grey colored MVB followed by fusion of the MVB with the plasma membrane resulting in release of the exosomes into the extracellular milieu. MVs are budded off the plasma membrane followed by release into the extracellular milieu.

promotes angiogenesis, suggesting that differential packaging of miRNA between normoxic and hypoxic cells can promote angiogenic activity in endothelial cells³⁹⁸. Similarly, exosomes secreted by MM cells chronically exposed to hypoxia are enriched in oncogenic *miR-135b*, which inhibits expression of FIH-1, an enzyme that inhibits HIF-1 α ³⁹⁹. In this way, exosomes from hypoxic MM can increase HIF-1 α activity and thus, increase vascular tube formation. EVs derived from CML cell lines can promote expression of angiogenic growth factors and cytokines, such as VEGF and IL8, through the activation of Src^{400,401}. Similarly, EVs released by the acute promyelocytic leukemia (APL) cell line NB4, representing a subtype of AML, are taken up by HUVEC cells where they promote increased survival and angiogenic function of HUVECs, also via stimulation of VEGF and IL-8 secretion⁴⁰². Interestingly, unlike with CML and BMSC³⁸¹, there was no transfer of the oncogenic fusion transcript, in this case PML–RARa⁴⁰².

EVs have been shown to alter migratory and adhesive potential in both endothelial and leukemic cells. Exosomes from MM-BMSC cause increased migration and adhesion of MM cells³⁹³. While the precise mechanism remains unclear, this effect is likely via CCL2 packaged in the exosomes³⁹³. Exosomes derived from the human CML cell line LAMA84 caused increased expression of the cell adhesion molecules ICAM-1 and VCAM-1 as well as IL-8 in HUVECs in a dose- and time-dependent manner, which increased the motility and stimulating tube formation of the HUVECs⁴⁰⁰. In contrast, long-term (12 - 24 h) exposure of HUVECs to CML-derived exosomes caused a reduction in CXCL12 and VCAM-1 mRNA, via transfer of *miR-126*⁴⁰³. The reduction in adhesive properties after longer exposures could potentially allow migration of CML cells past the endothelial layer in response to other chemokines. However, further evidence is required to determine if this is indeed the case.

1.4.8 EVs in resistance to cancer therapies

Resistance to chemotherapies can be attributed to a variety of molecular changes that can be affected by EVs⁴⁰⁴. For example, the overexpression of multi-drug resistant protein 1 (MRP-1) and/or P-glycoprotein (P-gp) has been proposed to contribute to resistance against chemotherapeutics in AML^{405,406}. MRP-1 expressing cell lines can transfer chemotherapeutic resistance *in vitro* via EVs, most likely via the direct transfer of MRP-1; however, the transfer of miRNA *miR-19b* and *miR-20a* is also suggested to contribute to the observed chemoresistance⁴⁰⁷. Multi-drug resistance (MDR), mediated by over-expression of P-gp encoded by the human MRD1 gene (ACBC1), can be transferred from resistant T-ALL cells to sensitive cells via microvesicles⁴⁰⁸. As described above, transfer of Gal-3 from BMSC can also promote chemoresistance of ALL to vincristine³⁸³. Thus, transfer of proteins and miRNA can transfer a resistant phenotype across a population of cells, which can increase the rate a patient acquires chemotherapeutic resistance. Rituximab is an anti-CD20 antibody used to treat some lymphomas and CLL by inducing destruction of CD20-expressing cells. Rituximab leads to cellular destruction through numerous mechanisms including direct induction of apoptosis, complement-dependent cytotoxicity (CDC), and antibody-dependent cytotoxicity (ADCC). Exosomes released from B cell lymphoma cells can express CD20 and treatment with Rituximab increases the number of EVs released^{409,410}. EVs expressing CD20 were found to effectively bind both the anti-CD20 Rituximab and complement, thus acting as a decoy to protect cells from CDC. Interestingly, inhibition of exosome release, by blocking the function of the ATP-binding cassette (ABC) transport protein ABCA3 restored Rituximab sensitivity⁴⁰⁹. This suggests that modulation of exosome levels could underlie the therapeutic resistance seen with ABCA3-expressing lymphomas⁴¹¹. It is possible that EV-

mediated sequestration of other monoclonal antibody-mediated therapies could reduce their efficacy, like the EV-mediated enhancement of immune evasion described above.

1.4.9 Clinical applications for EVs in HM

As mentioned, EVs have distinct protein and RNA profiles that differ significantly from healthy control EVs and reflect the cell of origin, as well as the ability to cross the BBB and the BCSFB^{412,413}. EVs released from cancer cells in any compartment in the body could likely be detected in peripheral blood. As such, EVs have been studied as early detection mechanism, for detection of MRD, for detection of relapse, and for monitoring treatment response. The abundance of cancer derived EVs may be of clinical benefit by overcoming the need to identify malignant cells for cytogenetic or flow cytometric analyses.

Increased EV plasma concentration was associated with more severe disease states in CLL, WM, and MM; whereas fewer EVs were found at later stages in patients with myelodysplastic syndromes⁴¹⁴. Although this initial study examined a small number of patients, other reports have also found an increased number of circulating EVs in patients with AML, CLL, or MM compared to healthy controls^{365,415}. For HL, there were higher levels of EVs in classical HL patients compared to healthy controls⁴¹⁶. Therefore, it is possible that total EV concentration could be used as an initial indicator of tumor burden, which may be of prognostic value particularly as a measure of the response of disease to treatment⁴¹⁷. Detecting tumor burden using EVs in peripheral blood could bypass the limitations and risks of invasive sampling of bone marrow or other tissues. It is possible that EVs could be identified even when the presence of malignant cells is below the detection threshold of currently used assays and histopathological techniques.

The potential diagnostic applications for EVs go well beyond concentration of circulating EVs as EVs can package actionable mRNA transcripts and/or proteins. This cell-specific packaging can

provide important diagnostic and functional information about the HM. For example, exosomes from AML can contain transcripts important both for prognosis (e.g. FLT3-ITD, NPM1) and for identifying targeted treatment (e.g. FLT3-ITD, IGF-IR, CXCR4)³⁷⁴. Microvesicles isolated from B cell precursor leukemia cell lines LAMA-87, REH, SHI-1, TOM1 and RS(4:11) carry the fusion transcript from their cells of origin³⁸¹. In diffuse large B cell lymphoma, specific mutations present in transcripts from four out of five different cell lines were also found in the EVs secreted from these cells⁴¹⁸. In CML, BCR-ABL fusion proteins and RNA transcripts could be detected in EVs, while MLL-AF6 and TEL-AML1 transcripts and proteins were detected in AML and B-ALL microvesicles, respectively³⁸¹. However, it is critical to determine in what HM or situations specific fusion transcripts are not packaged in EVs to limit this diagnostic strategy appropriately since false negatives will occur if fusion transcripts are not packaged into EVs.

Multiple miRNA species have been found to distinguish EVs in healthy individuals from those in cancer patients. Notably, increased levels of *miR-155* in EVs has been found to be associated with AML, CLL, and WM whereas EVs from patients with MDS and MM express low levels of *miR-155*^{365,419}. In MM, the miRNAs *let-7b* and *miR-18* are significantly associated with progression-free survival and overall survival, and are predictive of prognosis⁴²⁰. HL-related miRNAs present in circulating EVs include *miR-24*, *miR-155*, *miR-127*, *let7*, which could be used to monitor therapy response and relapse⁴¹⁶. Additionally, an EV-specific mRNA profile for CLL has been identified, which could be explored further for early diagnosis of CLL⁴²¹.

With respect to EV protein cargo, several tumor-associated proteins have been found in EVs. Initial studies revealed that signature proteins representing B cell neoplasms (CD19), MM (CD38), myeloid tumors, (CD13) and HL (CD30) are packaged into EVs⁴¹⁴. Classical Hodgkin's lymphoma (cHL) cell lines L540, L428, KM-H2, and L1236 release EVs that are positive for the

cHL markers CD30 and CD147⁴²². In addition, CD19⁺ and/or CD37⁺ EVs correlate with high tumor burden in CLL patients, suggesting that these particular protein biomarkers could be used to monitor treatment success in CLL^{365,423,424}. In MM, CD138⁺ microvesicles are associated with both disease phase and therapeutic response⁴²⁵. EVs may then provide a sensitive and non-invasive method, with quick turnaround, for characterizing HMs and identifying protein targets for therapies, such as Rituximab discussed above or Brentuximab vedotin, a monoclonal anti-CD30 drug conjugate⁴²⁶.

EVs have also been suggested to track some leukemia-associated complications. For example, the procoagulant activity of EVs has been linked to venous thrombosis and disseminated intravascular coagulant, which are known complications in patients with HM^{427,428}. Treatment of HUVECs with NB4-derived EVs caused an increase in the surface expression of tissue factor, which is not normally expressed by this cells⁴⁰². This observation may explain the increase in coagulation seen in APL. An initial study of 53 patients with acute leukemias such as B-ALL, T-ALL, and AML found that patients with thrombotic events had increased procoagulant activity of EVs measured *in vitro*⁴²⁷. In this small study, EV procoagulant activity was found to be a promising predictive measure for hemostatic complications in patients with acute leukemias, which was not predictable in this patient cohort based on standard laboratory information⁴²⁷.

The use of EVs for treatment of HM is also an intensive area of investigation with specific interests in EVs as drug/gene delivery vehicles, and for immunotherapy or as cancer vaccines. As drug/gene delivery vehicles, EVs are a promising tool as their cargo is naturally protected, their small size reduces clearance by phagocytes, they exhibit low immunogenicity if generated from autologous sources, they have inherent tissue specificity, and they can readily be incorporated into target cells⁴²⁹⁻⁴³¹. Loaded EVs can be generated by modification of the parent cell or by direct loading of the

EVs⁴³². A number of studies using this approach for solid tumors have been described (e.g.^{433,434}). Recently, this strategy has been applied to HM. Taking advantage of the fact that CML has a high expression of IL-13 receptor, HEK293T cell-derived exosomes were engineered to express recombinant IL-13, and loaded with BCR-ABL silencing RNA⁴³⁵. These EVs were able to reduce CML proliferation both *in vitro* and *in vivo*⁴³⁵. In addition, miR-328-loaded exosomes reduced imatinib-resistance in CML cells *in vitro*⁴³⁶. Liposomes generated with the same lipid composition as EVs have been loaded with recombinant Apoptosis ligand 2/TNF-related apoptosis-inducing ligand (Apo2/TRAIL) to increase the efficacy of the recombinant Apo2/TRAIL protein in targeting leukemia and lymphoma cell lines *in vitro*^{437,438}.

EVs carrying tumor antigens are able to activate an anti-tumor response and could have utility in immunotherapy or as cancer vaccines^{431,439}. Here, tumor-specific protein antigens or tumor-generated exosomes are incubated with dendritic cells (DC), which release exosomes that carry tumor antigens associated with MHC and co-stimulatory molecules. Injection of DC-derived exosomes (dexosomes), can in turn promote activation of cytotoxic T cells and a cytotoxic antibody response⁴⁴⁰⁻⁴⁴³. In this way, the immune system is primed to target the cancer cells. Some groups have examined these approaches in HM. An increased effect of anti-tumor toxicity was seen with DC loading of tumor-derived exosomes from leukemia cells *in vitro*^{444,445} and in a mouse model of leukemia⁴⁴⁵. In a small clinical trial of 15 MM patients, injection of DC-derived exosomes loaded with tumor peptides resulted in a clinical response in one-third of patients⁴⁴⁶.

For all these clinical applications, more extensive studies with larger patient cohorts and with different HM are needed to determine if any of these measures can be applied in patients or if clinically important benefits can be achieved. Regardless, the evidence to date suggests that

detection of disease specific EVs could be more sensitive and less invasive to inform treatment decisions, track treatment success, diagnose MRD, predict complications, and improve treatments.

1.4.10 Future Considerations

The idea that EVs from patients with HM may be used as a diagnostic, prognostic, or therapeutic tool is constantly evolving. Treatment monitoring using EVs from peripheral blood could occur more frequently and less invasively than biopsies of bone marrow, or lumbar punctures. Real-time assessment of disease status could be made possible using techniques that evaluate specific RNA and/or protein signatures of EVs. In addition, efficacy of therapy can be improved based on the detection of targetable proteins in EVs, such as CD20^{409,414}, PD-L1³⁶², or BCR-ABL1²⁴⁰. Moreover, targeted reduction of EV secretion may also restore immunotherapy efficacy^{370,409}. Another approach for EV based therapeutics is the inhibition of EV release to reverse EV-mediated alterations to the stromal cell environment and to suppress EV-mediated mechanisms for evading immunologic surveillance. This could subsequently promote normal immune cell function and development and inhibition of cancer cell growth.

However, the lack of in-depth understanding of EV release and uptake mechanisms makes it difficult to fully optimize strategies that target EVs directly. In addition, the side effects of inhibition of EV release on normal immune development or activation cannot be predicted, as the normal functions of EVs in immune system development and activation are not fully understood. Furthermore, direct targeting of endogenous EV transcripts could be even more challenging as there is no known method for selective manipulation of EV cargo without concurrent effects on bystander cells. For a more comprehensive analysis of the challenges in using EVs as therapeutics, we suggest the recent opinion article by Burnouf et al.⁴³².

In the short term, it is critical to determine the precise relationship of EVs (with reference both to their total concentration and to their biological characteristics) with disease stage, treatment response, relapse, and MRD in well controlled patient populations. Additionally, it is not fully clear if the levels of EVs in peripheral blood are a sensitive or specific enough measure for the detection of cells residing in extramedullary sanctuary sites, such as testes and CNS, which are protected from chemotherapeutics or targeted therapies ²²⁰.

EVs can affect normal cellular function as well as directly impact progression of HM through their ability to package a vast array of cargo. Harnessing the power of EVs will require a thorough detangling of the intricate threads linking EVs to specific functions and/or target cells. Nonetheless, regardless of the current limitations, the potential of EVs in improving diagnosis and treatment should encourage the rapid development and understanding of these small but powerful messages in the diagnosis, prognosis, and treatment of HM.

1.5 Research hypotheses and questions

Research hypotheses

1. Small RNA transcripts (especially miRNA) from plasma EVs could facilitate less invasive detection of MRD in pediatric B-ALL.
2. Pediatric B-ALL lymphoblasts package RNA transcripts into EVs using specific motifs and mechanisms.
3. Sample collection tube treatments alters EV physical properties and cargo.

Research questions

1. a. What is the miRNA signature of pediatric B-ALL?
b. What is the potential of blood EVs (physical characteristics and bioactive cargo) for monitoring pediatric B-ALL?

2. What is the cargo packaging mechanism of RNA into pediatric B-ALL EVs?
3. What are the effects of collection tube treatment on EV properties?

1.6 Publications arising from data presented in this thesis.

The results presented in this thesis have been published as follows:

1. Longjohn, M. N., Hudson, J. B. J., Smith, N.C., Rise, M., Moorehead, P., & Christian, S. L. (2020). Deciphering the messages carried by extracellular vehicles in hematological malignancies. *Blood Reviews*. doi: 10.1016/j.blre.2020.100734. This literature review is presented in chapter 1. The data presented in this chapter is used under copyright license from publisher Elsevier, with the copyright permission presented in appendix A.

Author contributions: MNL conceptualized and prepared the manuscript with contributions and edits by JBJH, MR, NCS, PM and SLC. All authors approved the final manuscript.

2. Longjohn, M. N., Squires, W. R. B., & Sherri L. Christian (2021). Meta-analysis of microRNA profiling data does not reveal a consensus signature for B cell acute lymphoblastic leukemia. *Gene*. This data is presented in chapter 3. The data presented in this chapter are used under the Creative Commons Attribution 4.0 international license as stipulated by Gene. The full license is presented in appendix A.

Author contributions: MNL and SLC conceived and designed the experiments and prepared the manuscript. MNL and WRBS did the data analysis. All authors approved the final manuscript.

3. Modeline N. Longjohn¹, Jo-Anna B.J. Hudson², Paul C. Moorehead², Lourdes Pena-Castillo^{3,4} Simi Chacko⁵, Stephen Lewis⁵, Sherri L. Christian^{1*} On behalf of the “PRrecision Oncology For Young peopLE (PROFYLE) program” (2023). Extracellular

vesicle small RNA cargo discriminates non-cancer donors from pediatric B cell acute lymphoblastic leukemia patients *In Revision*. This data is presented in chapter 4.

Author contributions: MNL conceived and designed the experiments, with contributions from JBJH and PCM and oversight from SLC. MNL did all experiments and data analyses under the supervision of SLC with input from LPC. SC and SL performed RNASeq. MNL prepared the manuscript and received inputs from all co-authors. All authors approved the final manuscript.

4. Modeline N. Longjohn¹, Jo-Anna B.J. Hudson², Paul C. Moorehead², Lourdes Pena-Castillo^{3,4} Simi Chacko⁵, Stephen Lewis^{5,6,7}, Sherri L. Christian^{1*} On behalf of the “PRrecision Oncology For Young peopLE (PROFYLE) program” (2023). Small RNA transcripts are selectively packaged into pediatric acute lymphoblastic leukemia-derived extracellular vesicles. *In Prep*. This data is presented in chapter 5.

Author contributions: MNL conceived and designed the experiments, with contributions from JBJH and PCM and oversight from SLC. MNL did all experiments and data analyses under the supervision of SLC with input from LPC. SC and SL performed RNASeq. MNL prepared the manuscript and received inputs from all co-authors. All authors approved the final manuscript.

Co-authorship statement

I, Modeline N Longjohn, in collaboration with my supervisor, Dr. Sherri Christian conceptualized and designed all experiments with contributions from Drs. Jo-Anna Hudson and Paul Moorehead. I performed, and analysed all experiments presented herein and wrote all portions of this thesis, unless specified below.

Key contributions to this thesis were made by collaborators as outlined below.

In chapter 1, Drs. Nicole Smith, and Jo-Anna Hudson contributed two sections each, while Drs. Matthew Rise, and Paul Moorehead made edits and suggestions.

In chapter 2, transmission electron microscopy was performed by Stephanie Tucker, while flow cytometry was performed with training from Dr. Nicole Smith, Stephanie Tucker, and Chris Corkum.

In chapter 3, Willow R. B. Squires performed analyses in cytoscape and generated cytoscape figures.

In chapter 4, Simi Chacko and Dr. Stephen Lewis sequenced the EV small RNA at the Atlantic cancer research institute, Moncton while Dr. L. Pena-Castillo facilitated the bioinformatics.

In chapter 5, Simi Chacko and Dr. Stephen Lewis sequenced the EV small RNA at the Atlantic cancer research institute, Moncton while Dr. L. Pena-Castillo facilitated the bioinformatics.

In chapter 6, Mohamad A. S. Khaled used NTA to quantify the EVs in plasma and urine, performed EV pulldown from plasma and urine and ran the Agilent bioanalyzer to quantify the RNA content of EVs.

Chapter 2: Methods

2.1 B-ALL Wet lab methods

The wet lab methods are summarized to give a clear understanding of the experiments (**Fig. 2.1**).

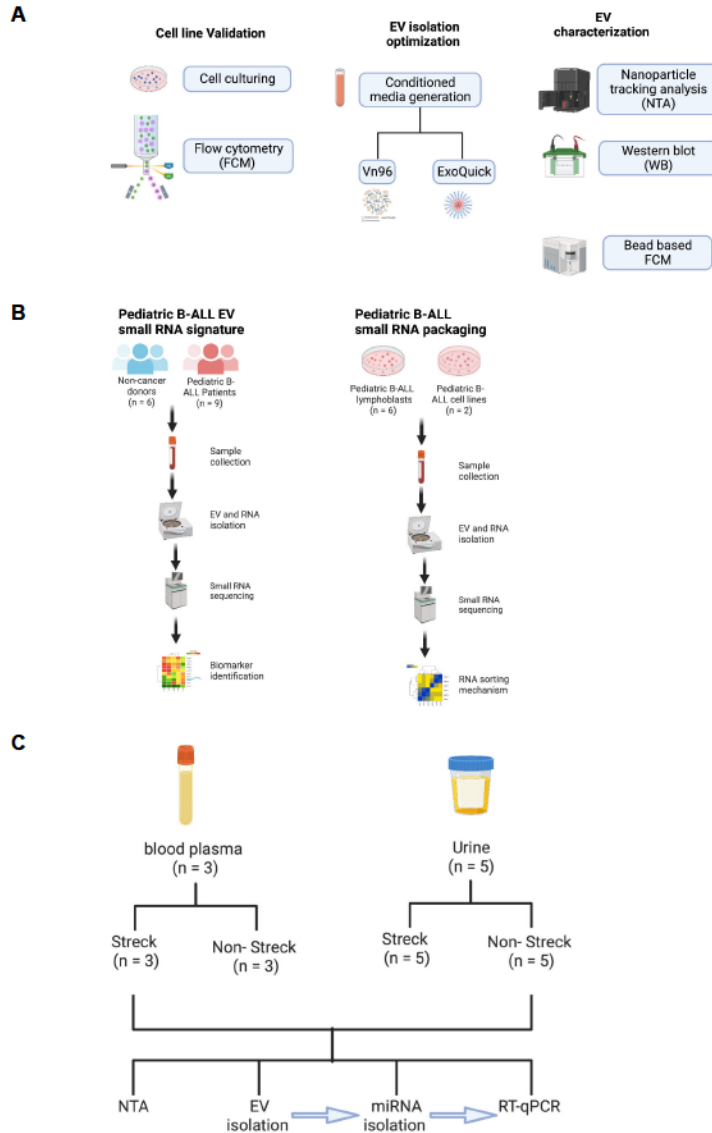


Fig. 2.1. Summary of wet lab experiments. A. cell line validation, EV isolation optimization and EV characterization techniques. B. Pediatric B-ALL EV small RNA signature and cargo packaging C. Effect of preservation on biofluid EVs.

2.1.1 Immortalized Cell lines

2.1.1.1 Validation of cell lines

The pediatric cell lines used as study models for this study are CCRF-SB, RCH-ACV and UoC-B1 and Burkitt's lymphoma cell line Ramos (**Table 2.1**).

Table 2.1. Cell line study models used for the study.

Cell line	Disease	Patient
CCRF-SB	B-ALL	11.5-year-old male patient with normal karyotype
RCH-ACV	Relapsed BM B-ALL	8-year-old female patient, Pre-B-ALL aka CALLA, hypodiploid karyotype with t(1;19), TCF3-PBX1 (E2A-PBX) fusion
UoC-B1	B-ALL	15.5-year-old female, with t(17;19) TCF3-HLF (E2A-HLF) fusion
Ramos	Burkitt's lymphoma	3-year-old male, with t(8;14) <i>c-myc/IgH</i> and TP53 mutations

2.1.1.2 Cell line tissue culture

CCRF-SB (ATCC no.: CCL-120, ATCC; Manassas, VA) was obtained from the American Type Culture collection (ATCC). RCH-ACV (DSMZ no.: ACC 548, DSMZ, Germany) was obtained from the DSMZ – German collection of microorganisms and cell culture, while UoC-B1 was a kind gift from the laboratory of Dr. Williams Evans from St. Jude's Children's Research Hospital, Memphis, Tennessee, USA. Ramos (ATCC no.: RA 1, ATCC; Manassas, VA) was obtained from ATCC. CCRF-SB, RCH-ACV and UoC-B1 were maintained in RPMI 1640 media (Gibco, Thermofisher, Cat. #: 11875093) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Thermofisher, Cat. #: 16000044) and 1% penicillin and streptomycin (5,000 u/mL, Thermofisher Cat. #: 15070063) to generate RPMI1640 complete media. Ramos cell line was

cultured in RPMI1640 complete media further supplemented with 1% sodium pyruvate and 0.1% β -mercaptoethanol. Cell lines were cultured in their respective complete media at 37 °C and 5% CO₂ until up to 70% confluency is reached.

2.1.1.3 Cell culture microphotography

Cells in suspension (for each cell line) are placed on a cover slip and visualized under a Leica DM IL ILED microscope, associated with the infinity camera and the Infinity Analyse software (**Fig. 2.2**).

2.1.1.4 Flow cytometry validation

Cells (from all four cell lines) were harvested from RPMI1640 complete media and washed in FACS buffer (1x Phosphate buffered saline – PBS supplemented with 1% heat-inactivated fetal bovine serum). Washed cells were then resuspended with FACS buffer and stained with 0.125 μ g of CD10-APC (17-0106-42, eBioscience), 0.125 μ g IgM-PE-Cy7 (314531, Biolegend), 0.5 μ g CD79a BV241 (562852, BD Biosciences), 0.25 μ g CD19-PE (12-0199-42, eBioscience), 0.25 μ g CD20-PECF594 (560962, BD Biosciences), 0.125 μ g Pax5-Alexa Fluor 488 (562816, BD Biosciences) and 0.5 μ g TdT-BV510 (565229, BD Biosciences), or with 0.25 μ g of CD24-PE (12-0242-82; eBioscience) for 30 mins at 4°C. This was followed by cells being washed with FACS buffer and analyzed by flow cytometry (Beckman Coulter cytoflex). In some cases, FC receptors were blocked using Seroblock (BUF070A, Bio-Rad), and cells were fixed using IC fixation buffer (00-8222-49, eBioscience). Cells were permeabilized using 1X permeabilization buffer (Saponin based) (00-8333-56). For analysis and gating strategies, negative (no-staining),

isotype and positive staining controls were used. For each analysis, at least 10,000 events were captured. Flow cytometry output is shown in **Fig. 2.3**, summarized in Table 2.2.

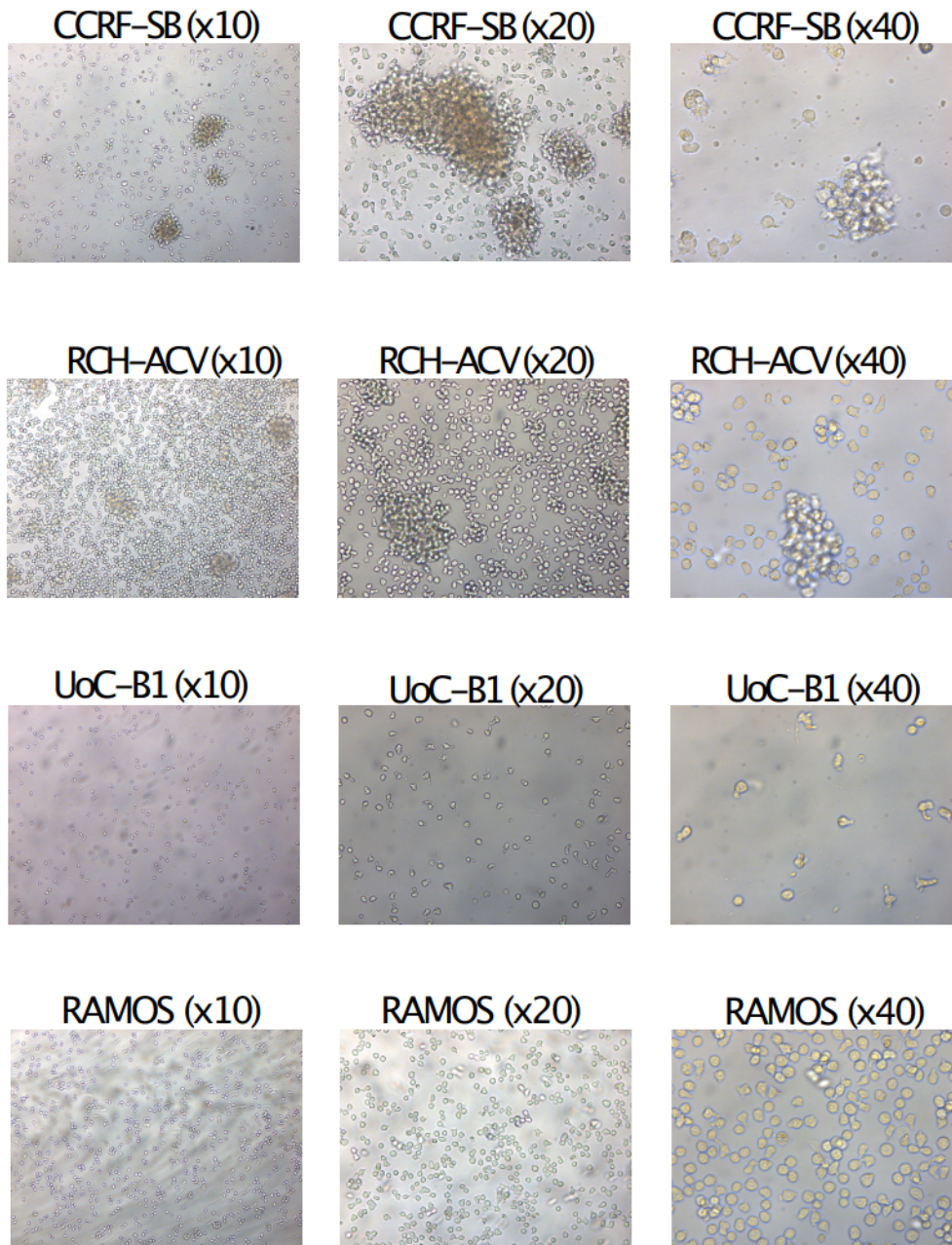


Fig. 2.2. Cell microphotography of cell line study models. Tissue culture dishes containing cultured cells are placed under microscope, and cell microphotography images were taken using microscope magnifications of x10, x20 and x40.

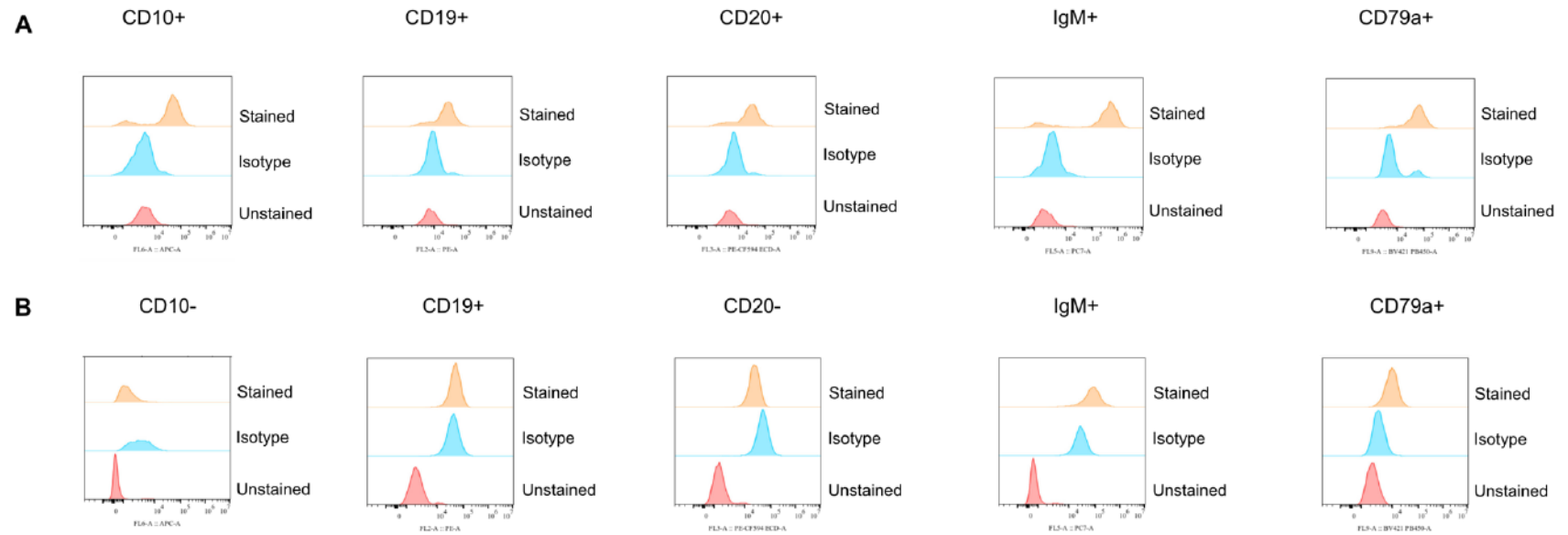


Fig. 2.3. Sample flow cytometry data showing expression of key surface and intracellular markers of B cells of different cell lines. (A) RAMOS (positive control) and (B) pediatric B-ALL (RCH-ACV) were validated using expression of surface markers (CD10, CD19, IgM, CD20) and intracellular marker (CD79a).

Marker	CD10-APC	CD19-PE	CD20-PE- CF594	CD79a- BV421	IgM-PE- Cy7	PAX5- Alexa Fluor 488	TdT-BV510
Marker location	Surface	Surface	Surface	Intracellular	Surface	Intracellular	Intracellular
Cytoflex channel Laser	APC 638nm	PE 488nm	ECD 488nm	PB450 405nm	PC7 488nm	FITC 488nm	KO525 405nm
Fixation Stability Cell Line	Stable	Not stable	Not stable	Stable	Stable	Stable	Stable
RAMOS	+	+	-	+	+	+	-
CCRF-SB	-	+	+	+	+	+(low)	-

RCH-ACV	-	+	-	+	+	+	+ (low)
UoC-B1	+	+	-	+	+	+	+ (low)

Table 2.2. Summary B cell immunophenotype of cell line study models.

2.1.2 Pediatric B-ALL primary lymphoblasts

2.1.2.1 Patient samples and characteristics

This study was approved by the Institutional Review Board of the Memorial University of Newfoundland and Labrador (HREB #2018.069) (Appendix B1). Human primary bone marrow lymphoblasts (BMMC) were isolated by Ficoll isolation from bone marrow aspirates collected in EDTA tubes obtained from the British Columbia Children's Hospital biobank, British Columbia (HREB # H13-03111). All methods were carried out according to the Canadian Research Tri-Council policy stipulations on ethical conduct for research involving humans (https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2018.html). Donors did not fast. See **Table 2.3** for donor information.

2.1.2.2 B-ALL lymphoblast tissue culture

B-ALL patient lymphoblasts were cultured as optimized by the Bruserud group with modifications⁴⁴⁷. Briefly, B-ALL patient lymphoblasts were revived from frozen vials and maintained in StemSpan™ serum-free expansion medium II SFEM II (Stemcell technologies, Cat. #: 09655) supplemented with human recombinant IL-3 (Stemcell technology, Cat # 78040.1), stem cell factor (Stemcell technology, Cat # 78062.1) and Flt3-ligand (Flt3/Flk-2) (Stemcell technology, Cat # 78009.1) and 1% penicillin and streptomycin (5,000 u/mL, Thermofisher Cat. #: 15070063) at 37 °C and 5% CO₂. Cell-conditioned media (CCM) was collected by centrifugation as appropriate for the different isolation methods outlined below.

Table 2.3. Characteristics of patient or non-cancer donor (NCD) blood plasma samples

Sample	Gender	Age at collection (years)	Donor diagnosis	PB WBC Count (x10 ⁹ /L)	Plasma volume (μl)	EV small RNA (ng)	Read count
NCD1	Male	2	Spinal muscular atrophy	NA	1000	76.64	1.25E+07
NCD2	Male	9	Acute tonsilitis	NA	1000	72.84	1.11E+07
NCD3	Male	3	Spinal muscular atrophy	NA	1000	29.89	2.03E+07
NCD4	Male	17	Epilepsy	NA	1000	18.89	2.35E+07
NCD5	Female	17	Epilepsy	NA	1000	1.73	1.54E+07
NCD6	Female	12	Spinal muscular atrophy	NA	1000	0.05	2.94E+07
2021PB07B_E	Male	11	ETV6::RUNX1 pediatric B-ALL	NA	1000	110.15	9.64E+06
2021PB08B_E	Male	9	ETV6::RUNX1 pediatric B-ALL	1.20	480	227.06	1.39E+07
2021PB09B_E	Male	3	ETV6::RUNX1 pediatric B-ALL	24.40	480	314.02	8.44E+06

2021PB10A_E	Female	3	ETV6::RUNX1 pediatric B-ALL	16.20	370	9.28	9.76E+06
2021PB12B_E	Male	2	ETV6::RUNX1 pediatric B-ALL	2.20	460	204.73	7.11E+06
2021PB13A_p	Female	12	BCR::ABL1 like pediatric B-ALL	507	290	457.22	1.93E+07
2021PB14B_p	Male	17	BCR::ABL1 pediatric B-ALL	3.10	440	38.25	8.15E+06
2021PB15A_p	Female	17	BCR::ABL1 pediatric B-ALL	49.90	320	1.56	1.16E+07

2.1.3 Human biofluid samples

2.1.3.1 Ethics Statement

The pediatric B-ALL EV small RNA study was approved by the Institutional Review Board of the Memorial University of Newfoundland and Labrador (HREB #2018.069) and the research ethics board affiliated with the province of British Columbia and University of British Columbia (HREB # H13-03111) (Appendix B1 and B2). Written informed consent was obtained from all participants before enrollment. All methods were carried out according to the Canadian Research Tri-Council policy on ethical conduct for research involving humans (https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2018.html).

2.1.3.2 Pediatric B-ALL and Non-cancer donor (NCD) Participants

Human plasma from whole blood collected in EDTA tubes was obtained from the British Columbia Children's Hospital biobank and the Janeway children's hospital, Newfoundland, and Labrador. Whole blood was processed within 1 h of collection by centrifugation at $5000 \times g$ for 15 min at 4°C to separate the plasma fraction. The plasma fractions were stored at -80°C until used. Before use, frozen plasma aliquots were slowly thawed on ice and treated for use as outlined in subsequent sections. All methods were carried out according to the stipulations of the Canadian Research Tri-Council policy on ethical conduct for research involving humans (https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2018.html). Donors did not fast. See **Table 2.3** for donor information.

2.1.3.3 Streck versus non-Streck sample collection and processing

Peripheral blood was collected at the University of British Columbia from three donors in two sets. The first set was collected in Streck coated collection containers while the second set was collected in non-Streck containers (EDTA coated). The blood was processed to deplete blood cells, and the plasma was stored at -80°C at the BC hospital biobank until transport to the Christian lab where plasma samples were stored at -80°C.

Urine was collected in three ways. Firstly, for urine EV collection optimization, a donor collected urine in sterile container at RT and stored at -80°C until use. For Streck versus non-Streck comparison, urine was collected from seven donors in two sets of containers – Streck coated containers and non-Streck vessels.

2.1.4 EV isolation optimization

2.1.4.1 Vesicle- free media generation

Two aliquots of RPMI1640 complete media were prepared as previously described²⁶⁵. Briefly, 20% heat-inactivated FBS (RPMI-20%) was centrifuged at $100,000 \times g$ for 18 h at 4 °C in an SW-28 rotor (Beckman Coulter, Brea, CA) to deplete FBS derived vesicles, filtered through a 0.22 µm filter and stored at 4 °C. FBS-free RPMI complete media was prepared using all ingredients except FBS. Vesicle-free media for culturing was prepared by mixing vesicle-free RPMI-20% and FBS-free RPMI-complete media in a 1:1 ratio.

2.1.4.2 Cell conditioned media collection

Cell conditioned media (CCM) was collected as follows. Cells (from each cell line) was harvested by centrifugation at 500 x g for 5 mins at RT and resuspended in vesicle-free media at a ratio of 1×10^6 cells /ml. Cells are cultured in vesicle free media for 24 hours, followed by cell pelleting (centrifugation at 500 x g for 5 mins at RT) and apoptotic bodies and debris pelleting (centrifugation at 2,000 x g for 5 mins at RT). CCM is used for EV isolation and characterization using appropriate methods.

2.1.4.3 Polyethylene glycol-based isolation (ExoQuick)

CCM was centrifuged at $3000 \times g$ for 15 minutes to remove cells and cell debris. CCM supernatant was then incubated with ExoQuick TC precipitation reagent (System Biosciences, Palo Alto, CA) and incubated in an upright position for at least 12 hours at 4 °C. Post incubation, ExoQuick-TC/CCM mixture was centrifuged at 1,500 x g for 30 mins at RT, followed by centrifugation at 1,500 x g for 5 mins to get rid of any residual supernatant. Microsphere beads were prepared by washing in resuspension buffer, before mixing with the beige/white EV pellet at RT on an inverting shaker for 15 minutes. Purified EVs are separated from the beads by centrifugation ($8000 \times g$ for 5 minutes at RT). The EVs were present in the supernatant, which could be stored on ice for immediate use, or at $-80^{\circ} C$ until ready for use.

2.1.4.4 Peptide affinity precipitation (Vn96)

EVs were isolated from CCM as previously optimized^{270,448}. CCM was incubated with 24 μ l Vn96 solution overnight at 4°C. For pediatric B-ALL signature study, 250 – 1000 μ l plasma was incubated with Vn96 as described above. For Streck/non-Streck comparison study, EVs were

isolated from 1ml plasma and 4ml urine samples respectively. Post incubation, EV were pelleted by centrifugation at 10,000 x g for 10 mins at 4 °C, followed by washing thrice with 0.1 µm filtered 1x PBS supplemented with PMSF. EV-Vn96 is stored as a pellet at -80°C until use.

2.1.4.5 Size exclusion chromatography (SEC)

CCM was centrifuged at 500 x g to pellet cells and 2,000 x g to pellet apoptotic vesicles and debris.

2.1.5.2 Western blot analysis

Cell pellets were lysed in Radioimmunoprecipitation assay (RIPA) lysis buffer (20 mM Tris, pH 8, 137 mM NaCl, 10% glycerol, 1% Triton X-100 and 2 mM EDTA) at a concentration of 1×10^6 cells/100 μ l 1x RIPA lysis buffer supplemented with 2 μ g/mL aprotinin (Sigma Aldrich, St. Louis, MO, USA), 1% phosphatase inhibitor cocktail (Sigma Aldrich, St. Louis, MO, USA), 1 mM sodium orthovanadate (New England Biolabs, Ipswich, MA, USA) and 1 mM PMSF (Sigma Aldrich, St. Louis, MO, USA). Cell lysis samples were then incubated on ice for 10 mins, followed by centrifugation at $17,000 \times g$ at 4°C for 10 min. The supernatant was collected and stored at -80°C until further use. Vn96-EV pellets isolated as described above were resuspended in 0.1 μ M filtered $1 \times$ PBS and mixed with 5x Laemmli sample buffer (300mM Tris buffer pH 6.8, 50% (v/v) glycerol, 25% β -mercaptoethanol, 0.05% (w/v) Bromophenol blue, 10% (w/v) sodium dodecyl sulphate) to a final concentration of 1x. Protein lysates were separated under reducing conditions for all markers except CD63. For all markers except Calnexin, 1 μ g for cell protein lysates were loaded, while 0.1 μ g was loaded for Calnexin. For EVs, proteins from 3×10^6 cell derived EVs were loaded. The proteins were separated using 10% and 12% SDS-PAGE as required. Proteins on gels were transferred onto nitrocellulose membranes, followed by blocking with 5% (w/v) skimmed milk in 0.1% TBST (17.53g NaCl, 4.85g Tris base, 0.1% Tween-20). Primary Abs were diluted in 5% (w/v) skimmed milk in TBST as follows: 1:1000 HSP90a/ β (Santa Cruz, sc-13119), 1:1000 CD81 (Santa Cruz, sc166029), 1:1000 HSC70 (Santa Cruz, sc-7298), 1:500 CD63 (Santa Cruz, sc-5275), 1:500 Albumin (Cell Signaling Technology- cat no.4929), 1:10,000 Apolipoprotein B (Abcam, ab139401), 1:1000 Apolipoprotein A (Santa Cruz sc-376818), 1:1000 Calnexin (Abcam, ab22595), 1:1000 BLNK (Cell signaling technologies, 12168), 1:1000 CD41 (Cell signaling technologies, CST 138075), and 1:1000 CD235 (Santa Cruz, sc-59182). HSP90,

HSC70, ApoA, and CD63 were detected using HRP-conjugated goat anti-mouse IgG (Santa Cruz-13119) while ApoB, CD41, Calnexin and Albumin were detected using HRP-conjugated mouse anti-rabbit IgG (Santa Cruz, sc-2357). CD235 was detected using goat anti-rat IgG-HRP HRP conjugated (Santa Cruz, sc-2006). All secondary Abs were diluted 1:2000 in 5% (w/v) skimmed milk in TBST. Western chemiluminescent HRP substrate (Immobilon ECL Ultra Western HRP Substrate) was used for detection. Western blot image acquisition was done on a gel documentation system (Bio-Rad, Hercules, CA), followed by simple image manipulation involving only brightness and contrast adjustments of the entire image.

2.1.5.3 Transmission electron microscopy

For pediatric B-ALL cell lines, EV-Vn96 Pellets were resuspended in 0.1- μ m-filtered 1x PBS and EVs dispersed from Vn96 peptide by digestion overnight with 25 μ g proteinase K enzyme (Sigma-Aldrich)²⁷⁰ at 37 °C. Post digestion, Vn96-EV-proteinase-K mixture was centrifuged at 17,000 \times g for 15 min to remove undigested material. For patient lymphoblast CCM, 5 μ l of CCM was used for TEM. Diluted dispersed EVs were placed on formvar-carbon electron microscope grids (Electron Microscopy services, Hatfield, PA, USA) and allowed to dry for up to 60 mins. Grids were floated sample-side down in pyrogen-free water followed by fixation with 3.7% paraformaldehyde for 15 min and two washes with pyrogen-free water by slow dropwise application for 60 sec each time. Grids were then contrasted with 2% uranyl acetate (w/v) for 6 mins, followed by one additional water wash as above. All solutions were filtered using 0.1- μ m syringe filters (4611; Pall Corp; Port Washington, NY). Dried grids were then viewed using a Tecnai Spirit transmission Electron Microscope (TEM) operating at 80KV (FEI; Hillsboro, OR).

2.1.6 RNA isolation and processing

2.1.6.1 Total small RNA isolation

mirVana miRNA isolation kit (Ambion, Life Technologies, Carlsbad, CA, USA) was used for total small RNA extraction following the manufacturer's instructions. EV pellet was resuspended in elution solution followed by stepwise isolation. Finally, 100 µl RNase-free water heated to 95 °C in a heat block was used to elute the total small RNA. Total small RNA in water was stored at -80° for later use. Total small RNA in water was concentrated using the Savant speed vacuum centrifuge concentrator (Thermo Scientific, Asheville, NC) at 43°C for 2 rounds of 30 mins each. 1 µl RNA was then used for quantification using the Agilent bioanalyzer (Santa Clara, CA). Concentrated total small RNA samples were then sent to the Atlantic Cancer Research Institute (Moncton, New Brunswick, Canada) for library preparation and sequencing.

2.1.6.2 RNA QC

The quality and quantity of the RNA were assessed using the Agilent small RNA kit (Agilent, 5067-1548, 5067-1550 and 5067-1549) and the HS RNA assay on the fragment Analyzer (Agilent, 5067-5576 and 5067-5577) for integrity and size distribution respectively.

2.1.6.3 Library Preparation

The library was prepared using 10ng RNA input or less and the NEXTFLEX Small RNA-Seq Kit v3 (Perkin Elmer, NOVA-5132-24) following the manufacturer's instructions. First, QIAseq miRNA Library QC Spike-Ins (Qiagen, 331535) was diluted 1/2 and then added to the samples resulting in a final dilution of 1/42. NEXTFLEX 3' 4N adenylated adapters were ligated to samples using a 1/4 adapter dilution and a 16°C overnight incubation. The adaptor-ligated samples were

purified with NEXTFLEX cleanup beads, and excess adapters were inactivated. Next, NEXTFLEX 5' 4N adapters were added to samples using a 1/4 adapter dilution followed by reverse transcription first strand synthesis. Reverse transcribed samples were subjected to a bead clean-up without any size selection. For the PCR enrichment, a different unique dual index (Perkin Elmer, NOVA-5132-24) was added to each sample, and 22 PCR cycles were used. Amplified libraries were purified without any size selection initially; however, due to high molecular weight templates, a final size selection was performed to obtain the final library. The quality of the library was assessed as follows; size distribution was determined with the D1000 assay on the TapeStation (Agilent, 5067-5582 and 5067-5583). The KAPA Library Quantification Kit (Roche, KK4824) was used to evaluate the concentration.

2.1.6.4 Small RNA Sequencing

Equimolar amounts of libraries were first sequenced on the iSeq 100 instrument (Illumina) using single-end sequencing of 1X100 to assess both library and pooling qualities. Libraries' inputs were rebalanced following Illumina's recommendations to ensure an equal representation of each sample. Libraries were then sequenced using the Novaseq 6000 instrument (Illumina). Samples were loaded on an S1 flow cell, and a single-end sequencing of 1X101 was used.

2.1.6.5 cDNA synthesis

cDNA synthesis was performed using the micro script miRNA cDNA synthesis kit (54415, Norgen Biotek) as described in their datasheet. Briefly, for each sample, 2x reaction mix, microscript microRNA enzyme mix, RNA template (between 2.5 – 4 ng, depending on the sample) were added to a single nuclease free PCR tube. Sample preparation was done on ice, and then transferred to a

thermocycler and cDNA synthesized using the following settings (37 °C for 30 minutes, 50 °C for 30 minutes, 70 °C for 15 minutes followed by a hold at 4 °C). Synthesized cDNA was then transferred to storage at -20 °C until use.

2.1.6.6 qPCR

qPCR was performed using Agilent miRNA qPCR mastermix (600583, Agilent technologies)

190 results respectively (Fig. 2.5A). Using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria⁴⁴⁹, results were filtered down to 31 and 24 results, respectively (Fig. 2.5A). Final inclusion datasets from GEO and EBI were 5 and 4, with 4 datasets overlapping in both databases (Fig. 2.5A). For meta-analysis of previously published DE miRNAs, literature with miRNAs in B-ALL were searched for in National Center for Biotechnology Information (NCBI) <https://www.ncbi.nlm.nih.gov/gds/> using the phrase ‘miRNAs in B cell acute lymphoblastic leukemia’. Datasets for re-analyses and meta-analysis were included based on presence of at least 3 healthy controls (HC), and over 10 patient B-ALL samples. Datasets where the patients were treated with chemotherapeutic drugs were excluded. Studies without HC were also excluded.

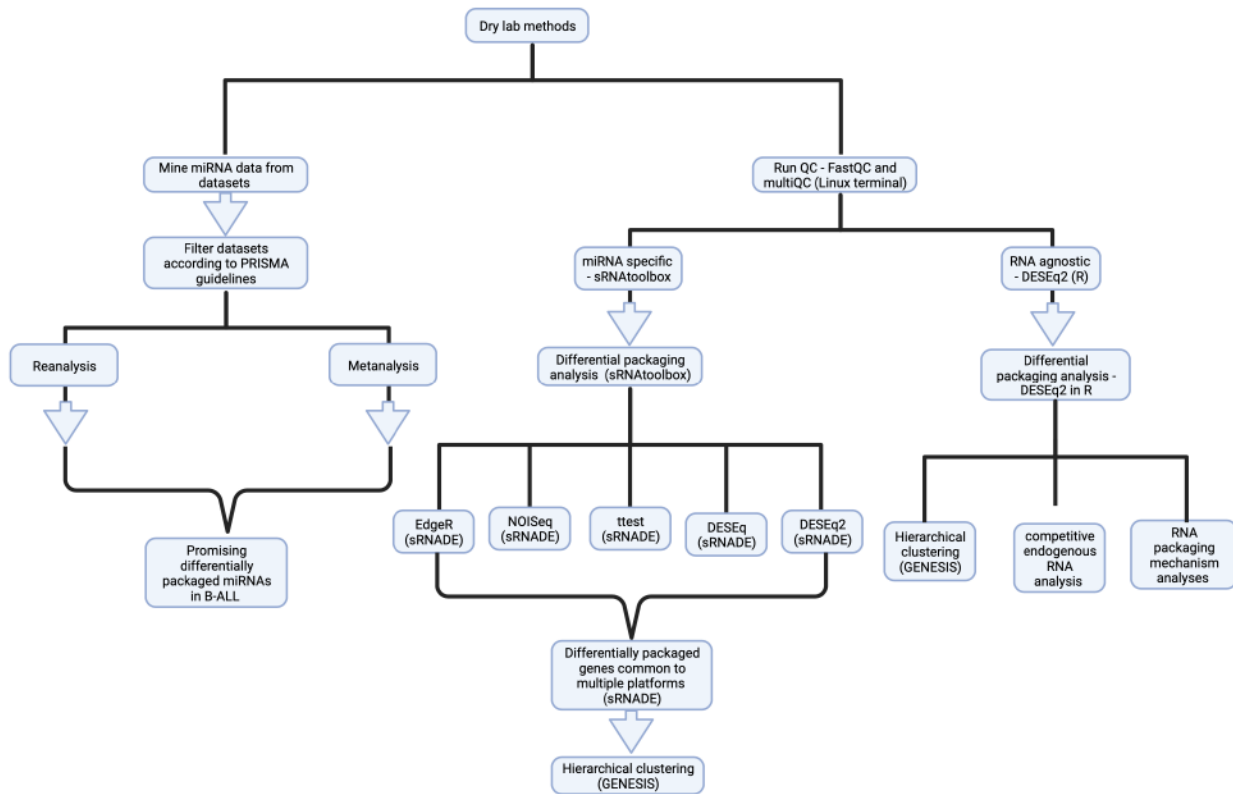


Fig. 2.4. Experimental design for dry lab experiments. Summary steps miRNA bioinformatics-based study (left) and detailed steps for RNASeq data processing, using the miRNA-specific platform- SRNAtoolbox (middle) for a miRNA-specific signature and the RNA agnostic platform – DESeq2 for RNA agnostic signature (right).

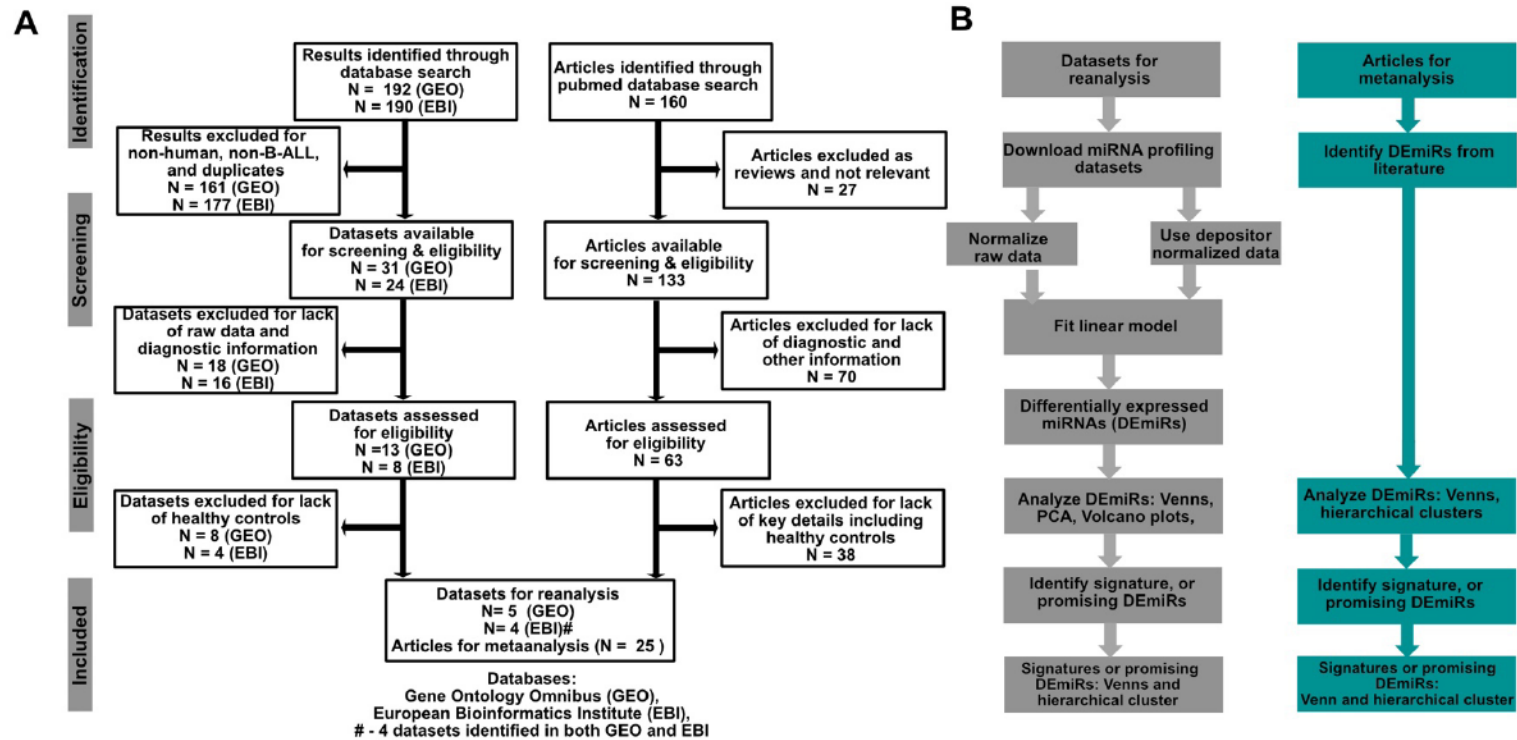


Fig. 2.5. Experimental design for bioinformatics study, and main steps for study identification and inclusion. For study identification, dataset and literature search was done in Gene expression omnibus (GEO) and PubMed databases. Using multiple rounds of exclusions, search results (for reanalysis – left and metaanalyses – right) were narrowed to the final dataset and literature samples. (B)

After data collection, raw data (microarray and PCR profiled) data was analyzed to identify differentially expressed miRNA (DEmiRs) between healthy control and B-ALL.

2.2.1.2 Re-analyses of raw data

The five microarray and PCR datasets analyzed are described in **Supp. file 3.1** and analyzed as shown in **Fig. 2.5B**. Each dataset contained healthy and B-ALL bone marrow and/or peripheral blood samples from pediatric and adult patients. Raw data was normalized using the methods stated in **Supp file 3.1** using the R statistical package (v 3.6.3). Normalized expression data was then fit to a linear model using Bioconductor tool LIMMA (v 3.28.14). Analysis of differential expression data between different groups (HC versus B-ALL; HC versus B-ALL subtypes) were performed, with a false discovery rate (FDR) of 5%. Venn diagrams were created with <http://bioinformatics.psb.ugent.be/webtools/Venn/>, Van de Peer Lab, Ghent University). Principal component analysis (PCA) was calculated using ‘ggplot2’ package (v. 3.3.3). Z scores for total miRNAs per dataset were calculated using the Biobase package (v 3.12). Hierarchical cluster analysis using median linkage was constructed using GENESIS software (v 1.8.1)⁴⁵⁰.

2.2.1.3 Meta-analysis of DE miRNAs

Differentially expressed miRNAs (DE miRNAs) from 25 papers (**Supp file 3.2, Fig 2.5A**) profiled by cDNA microarray, RNA-seq or PCR were manually curated into Excel. DE miRNA data were then analyzed by Venn diagrams as well as cluster analysis using GENESIS software (**Fig.2.5 B**).

2.2.1.4 *In Silico* Target prediction and Gene Pathway Enrichment Analysis

The top 8 promising miRNA names and mature sequences were verified using miRbase (<http://www.mirbase.org>, Manchester University, release 22.1). The putative target genes of the 8 promising miRNAs were predicted using miRwalk 3.0⁴⁵¹ and miRDB 6.0⁴⁵². The prediction parameters used perfect seed complementarity and high base-pair stability, to minimize false positives. Total targets identified from miRwalk and miRDB for each miRNA are shown in **Supp.**

file 3.3. The top 25 mRNA targets common to both analyses are shown in **Supp file 3.4** and were chosen for further analyses.

Gene pathway analysis was carried out using ClueGo plugin for Cytoscape software (v 3.8.2 National Resources for Network Biology)⁴⁵³ to identify significantly enriched biological processes (BP) and Kyoto encyclopedia of genes and genomes (KEGG), Reactome and wiki pathways.

2.2.2 *Pediatric B-ALL EV Small RNA*

2.2.2.1 Data Processing

Raw reads from RNASeq were processed using Illumina's basespace sequence downloader hub (v 2.2.1.6, <https://basespace.illumina.com/dashboard>, accessed on April 22, 2022), which runs on the Illumina sequence hub platform 2017. Next, adapter sequences were removed and size filtered using Cutadapt (v4.0)⁴⁵⁴, followed by QC using FastQC (v0.11.9)⁴⁵⁵ and MultiQC (v1.13.10)⁴⁵⁶. Alignment to a reference genome was conducted with Bowtie 2 (v2.4.5)⁴⁵⁷, followed by indexing using Samtools (v1.16.1)⁴⁵⁸ and viewing by Integrative genome viewer (v2.13.0)⁴⁵⁹.

Using feature count data, a stringent filter criterion was applied as follows – samples with reads in 8/8 pediatric B-ALL samples, and samples with reads in 6/6 NCD samples were delineated. Based on this stringent criterion, the list of transcripts was extracted and ran through the Ghent Venn diagram packaged <https://bioinformatics.psb.ugent.be/webtools/Venn/>.

A miRNA-specific B-ALL EV signature (significantly differentially packaged miRNAs (DPMiRs) between NCD and B-ALL) was identified using sRNAtoolbox⁴⁶⁰ (accessed December 2022). Within the sRNAtoolbox analysis, five pipelines were used: DESeq, DESeq2, NOISeq, EdgeR and T-test. For the RNA agnostic analysis of differentially packaged RNAs (DPRNAs),

read counts were obtained with featureCounts (v2.04)⁴⁶¹ in RStudio (v2022.02.3). R statistical environment (v3.4.1) was used to calculate the difference between the normalized read counts of RNA with the “relative log expression” (RLE) normalization method of Bioconductor package DESeq2 (v1.36.0)⁴⁶². Statistically significant differential packaging (DP) was determined using the false discovery rate correction (FDR) at a cut-off of 0.05 or lower, as indicated. The RNASeq data have been deposited in the NCBI Gene Expression Omnibus (GEO) under the accession number GSE239461 and GSE239467.

To identify exclusive transcripts, feature count data, a stringent criterion was used where transcripts with raw counts present in 8/8 cells (primary lymphoblasts and immortalized cell lines) or in 8/8 EVs were extracted. A Venn using the Ghent Venn diagram tool (<https://bioinformatics.psb.ugent.be/webtools/Venn/>) was carried out, to identify cell and EV exclusive transcripts. Pairwise and multi-group differential expression analyses were performed using the R statistical environment (v3.4.1). Principal component analysis (PCA), and MA plots were generated using DESeq2 (v3.16) and ggplot2 (v3.4.0). Unsupervised hierarchical clustering was performed on z scores using the GENESIS software (v1.8.1)⁴⁶³.

To identify the pattern of genes linked to the packaged genes and associated themes, gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) were analyzed using Cytoscape⁴⁶⁴ and R Bioconductor packages (clusterProfiler⁴⁶⁵, tidygraph, tweenr, pathview⁴⁶⁶, enrichplot and ggplot2).

To compare the abundance of tRNA-anticodons to the total tRNAs in the genome, the list of packaged tRNA genes were annotated and compared to the genome tRNA list obtained from the Lowe lab’s tRNA database. This was accessed from

<http://gtrnadb.ucsc.edu/genomes/eukaryota/Hsapi19/Hsapi19-pseudo-gene-list.html> on June 30, 2023, with the total list downloaded.

2.2.2.2 Post sequencing analysis

RNA transcripts from the RNA agnostic pipeline were using Ensembl Biomart⁴⁶⁷ into coding RNA (mRNA) and non-coding RNA (miRNA, tRNA, lncRNA, snoRNA, snRNA) and unclassified non-coding RNAs using NCBI (<https://www.ncbi.nlm.nih.gov> accessed on December 2, 2022), HUGO gene nomenclature committee (<https://www.genenames.org> accessed on December 2, 2022) and RNAcentral (<https://rnacentral.org/> accessed on December 2, 2022).

Pairwise and multi-group differential expression analyses were performed using the R statistical environment (v3.4.1). For RNASeq data exploration and quality check, plot counts, principal component analysis (PCA), and volcano plots were generated using DESeq2 (v3.16) and ggplot2 (v3.4.0). To identify the pattern of genes linked to the identified DPRNAs and exclusively packaged genes, gene set enrichment analysis (GSEA) and pathway analyses were performed using clusterprofiler (v4.6.2)⁴⁶⁵, tidygraph (v1.2.2), tweenr (v2.0.2), pathview (v1.46.1)⁴⁶⁸, enrichplot (v1.18.0) and ggplot2(v3.4.0). Unsupervised hierarchical clustering was performed on z scores of read counts that were regularized logarithmic transformation (rlog) by the DESeq2 platform, using the GENESIS software (v1.8.1)⁴⁶³. Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) analyses were carried out using Cytoscape⁴⁶⁴.

To identify the potential competitive endogenous RNA networks, the miRNA:mRNA, miRNA:lncRNA and miRNA:lncRNA interactions of annotated RNA subtypes were identified using miRcode (<http://www.mircode.org> accessed on December 9, 2022), miRwalk (<http://mirwalk.umm.uni-heidelberg.de> accessed on December 9, 2022), targetscan (https://www.targetscan.org/vert_80/ accessed on December 9, 2022), lncRRIssearch

(<http://rtools.cbrc.jp/LncRRlsearch/> accessed on December 12-18, 2022). ceRNA networks were constructed using Cytoscape⁴⁶⁴, and the connecting RNA extracted into a table.

2.2.2.3 In-silico target gene prediction

Target gene predictions for DPmiRs were carried out using target gene prediction software miRDB version 6.0^{452,469} and Target scan version 8.0⁴⁷⁰. miRDB predicts targets based on the miRDB MiRTarget algorithm, which predicts targets based on 3'UTR and seed sequence match. Predicted targets were ranked, including using target prediction scores. The predicted targets between both software were compared, and only those identified using both were taken as targets.

2.2.2.4 RNA packaging motifs

RNA motifs overrepresented in EV enriched, cell enriched, EV exclusive and cell exclusive signatures were identified using MEME Suite (version 5.5.1)⁴⁷¹. To identify motif site distribution, an 'any number of repeats' (anr) model was used to identify motif site distribution RNAs, with a motif count of 50. An E-value cut-off of <0.05 was used for determining significance.

Chapter 3: Meta-analysis of microRNA profiling data does not reveal a consensus signature for B cell acute lymphoblastic leukemia.

3.1 Introduction

Globally, acute lymphoblastic leukemia (ALL) is the most prevalent pediatric cancer and is the main cause of childhood cancer associated deaths⁴⁷². ALL is a clonal disease of B and T lymphocytes (B and T cells), characterized by arrest in cell development and accumulation of lymphoblasts¹⁵⁴. Typically, 80% of ALL arises in immature B lymphocytes (B-ALL), while the remainder arise in immature T cells (T-ALL) or other stages of B cell development⁴⁷³. ALL symptoms tend to be nonspecific, thus a multi-pronged approach that combines cytogenetic analysis, flow cytometry based immunophenotyping and polymerase chain reaction is taken for diagnostics¹⁵⁴.

In B-ALL, non-functional lymphoblasts accumulate in the bone marrow and eventually migrate into blood circulation where they crowd out other cells. Classification of B-ALL is based on the expression of B cell markers, such as CD10, CD19, CD20, CD22, CD34, HLA-DR, and CD79a, that are linked to the stage of B cell development when growth arrest occurs¹⁵⁸. Thus, based on cellular markers, B-ALL is subtyped into pro-B-ALL, early pre-B-ALL, pre-B-ALL (common acute lymphoblastic leukemia [CALLA] and non-CALLA) and mature B-ALL. Other parameters for classification include age of patient, initial leukocyte count and cytogenetics²¹².

The goal of this study is to identify consensus miRNA transcripts that differ between healthy and B-ALL patients. Because of the important role of miRNAs in regulating gene expression, a consensus miRNA signature of B-ALL could drive the field forward towards harnessing miRNAs for B-ALL diagnostics and therapeutics. Consensus miRNA signatures have been identified in

other diseases such as miR-331 and miR-195 distinguishes local and metastatic breast cancer⁴⁷⁴, miR-15b, miR-27b, miR-32, miR-130a amongst others in cervical cancer⁴⁷⁵, miR-15a-5p in CLL⁴⁷⁶ and miR-29b-1/miR-29a in AML. Some of these miRNAs have now been tested in clinical trials to gauge their feasibility for disease monitoring. Therefore, finding antisense miRNAs that distinguish pediatric B-ALL from controls is an important question to answer. To identify a signature of differentially expressed miRNAs (DEmiRs) in B-ALL, we identified datasets that we used for reanalysis of the raw data as well as identifying publications used for meta-analysis. Surprisingly, we found that there are no consensus DEmiRs between healthy control (HC) and B-ALL profiled, thus far. However, we identified 8 promising DEmiRs from the meta-analysis found to change in the same direction in multiple platforms and studies, which are known to be important in B cell development. These 8 DEmiRs could potentially form the basis for additional studies on miRNA in B-ALL.

3.2 Results

3.2.1 Comparison of healthy versus disease miRNA profiles revealed miRNA expression patterns.

The objective of this study was to identify a putative miRNA signature of B-ALL, which could be validated as a biomarker for better tracking of disease progression and response to therapy. To do this, we performed a two-pronged systematic analysis of the miRNA landscape in B-ALL: a reanalysis of B-ALL miRNA profile raw data and a meta-analysis of DEmiRs identified in the literature.

For the reanalysis experiments, five datasets (GSE31376, GSE23024, GSE2564, GSE59199 and GSE56484) passed the exclusion and inclusion criteria. GSE31376 and GSE23024 contained samples from different B-ALL subtypes (subtyped samples) while the rest did not contain subtype information. Of all 5 samples, GSE59199 and GSE56484 contained the fewest number of patient samples. Based on the format of the data and platform, each dataset was normalized as follows: GSE31376, GSE23024 and GSE2564 were normalized using quantile normalization, GSE59199 was normalized using print tip loess and scale normalization, and GSE56484 was normalized using between array and quantile normalization (**Supp fig. 3.1**). Boxplots after normalization of each dataset is shown in **Supp fig. 3.2**. Volcano plots which display unstandardized signal as log₂fold-change (logFC) against noise adjusted/standardized signal (i.e., p-value) are shown in **Fig. 3.1A**.

Further, principal component analysis (PCA) plots are shown in **Fig. 3.1B**. Although the normalization was successful, the number of genes that passed the statistical and logFC thresholds were variable between datasets. In addition, PCA analysis revealed that only 3 datasets (GSE23024, GSE2564 and GSE59199) resolved HC from B-ALL (**Fig 3.1B**).

Furthermore, DE miRNAs analysis was done in two sets. The first set included all datasets and compared HC to unsubtyped B-ALL and identified DE miRNAs within each dataset (**Fig. 3.2A**). There were no DE miRNAs common to all 5 non-subtyped datasets, however 8 DE miRNAs were common to 3/5 datasets (GSE23024, GSE31376 and GSE2564) (**Supp fig, 3.6**). Hierarchical cluster analysis of Z-scores was then performed to see if any expression patterns became apparent (**Fig. 3.2B**). However, the hierarchical cluster of miRNAs across all five datasets showed no clear signature with respect to disease state. Instead, the samples mainly clustered by dataset as can be seen by the color-coded

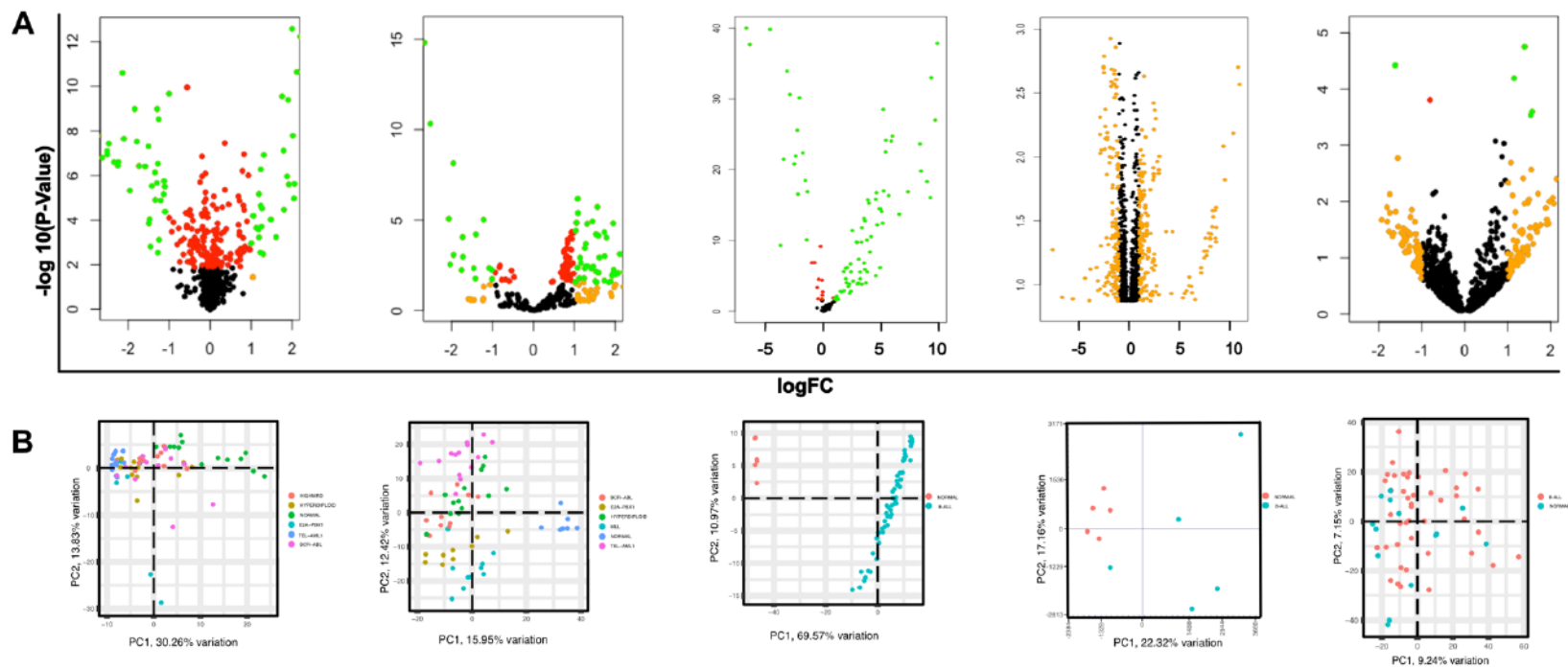


Fig. 3.1. Differentially expressed miRNAs (DEmiRs) re-analysis.

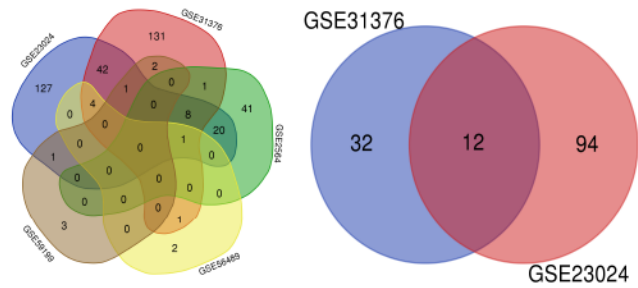
(A) Within each dataset, statistical significance (LogP values) of samples versus fold change (FC) were plotted on Volcano plots. (B) For each dataset, principal component analyses were generated. The samples were color-coded by sample type (healthy versus B-ALL). Note that the color-coding is dataset specific.

sample names, which grouped together, while HC grouped separately from B-ALL as shown by the colored bar. The second analysis included only the two datasets where B-ALL was subtyped. For these datasets (GSE31376 and GSE23024), each specific subtype was compared to HC (Hyperdiploidy, TCF3::PBX1, BCR::ABL, ETV6::RUNX1 (TEL::AML1), High Minimal residual disease-MRD and KMT2A rearranged) and DE miRNAs common to both comparisons were identified (**Fig. 3.2A**). Here, we found 12 DE miRNAs common to both datasets. The hierarchical cluster based on Z-scores of these 12 miRNAs (miR-181b/c/d, miR-27a, miR-34a, miR-130b, miR-134, miR-143, miR-145, miR-223, miR-593 and miR-650) are shown in **Fig. 3.2C**. Again, the data clustered by dataset and not by disease state as seen by comparing the color-coded sample names (datasets) to the color-coded bar (disease state). Next, we analyzed the logFC of the 12 DE miRNAs to compare the change in expression found within each dataset. We found that there was a conflicting change in the direction of expression in 10 miRNAs between the datasets where a gene that was increased in one dataset was decreased in the other dataset (**Fig. 3.2D**). Two miRNAs (miR-593 and miR-650) showed an increase in both datasets; however, they did not agree in magnitude of change. Overall, we were unable to define a clear signature that differentiated HC from B-ALL with the reanalysis of these datasets.

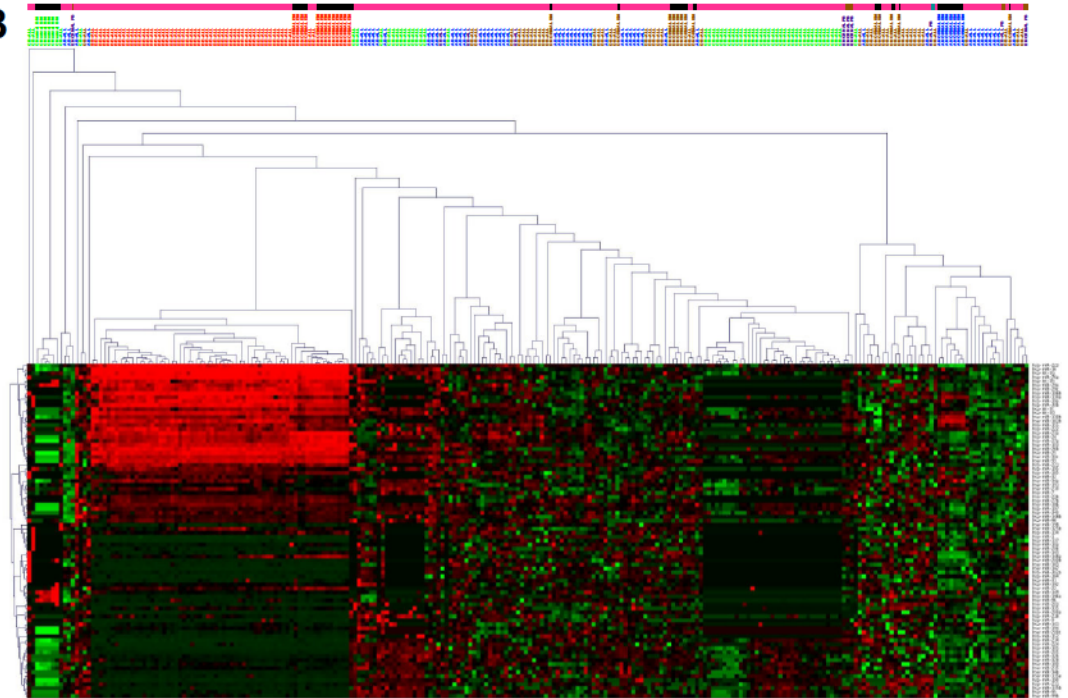
3.2.2 Meta-analysis of DE miRNAs show conflicting expression patterns for some miRNAs.

Twenty-five published studies were mined for DE miRNA data (Supplementary file 3.1). We found that miRNA profiling was done using three main methods: cDNA microarrays (32.5%), RNA sequencing (RNA-Seq) (7.5%) and polymerase chain reaction (PCR) (60%). Within platform

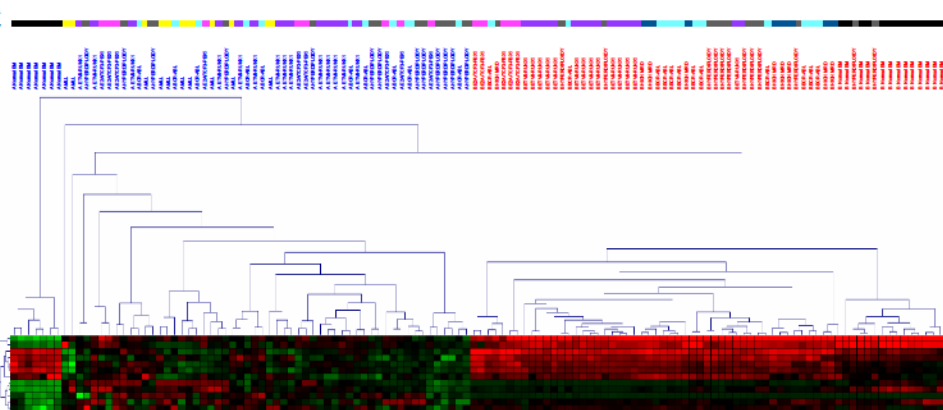
A



B



C



D

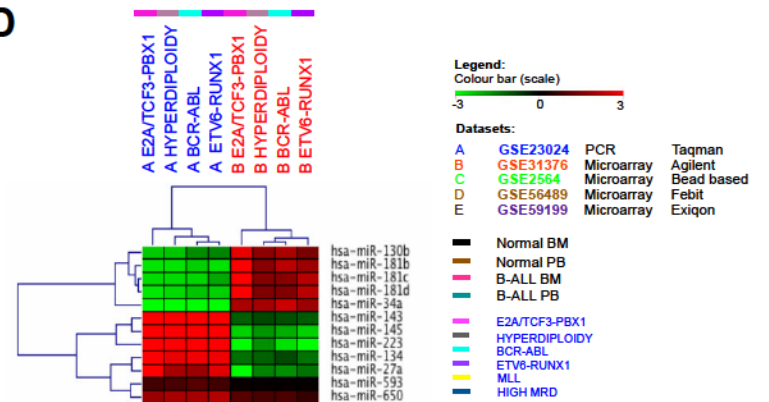


Fig. 3.2. Reanalysis identified overlapping miRNAs and DEmiRs (between datasets) show no clear consensus signature.

(A) DEmiRs for five datasets were analyzed by Venn diagram, with DEmiRs for non-subtyped datasets in a five-way Venn and DEmiRs for subtyped datasets in a two-way Venn. (B) Overlapping miRNAs across all five datasets were identified and expression data converted to Z-scores. Z-Scores of overlapping miRNAs were analyzed by unsupervised hierarchical clustering. (C) For subtyped datasets (GSE31376 and GSE23024), Venn diagram analysis (see panel A above) identified 12 overlapping miRNAs, which were analyzed by unsupervised average linkage hierarchical cluster analysis. (D) Fold change (FC) of the 12 DEmiRs overlapping between two subtyped datasets were analyzed by unsupervised average linkage hierarchical cluster analysis.

comparison between upregulated and downregulated DEmiRs showed that 12 DEmiRs overlapped with miRNA changing in both directions for microarray, 0 for RNASeq and 24 for PCR (**Fig. 3.3A**). Across platform comparison of DEmiRs upregulated in each of the 3 platforms showed an overlap of 28 DEmiRs that were upregulated in both microarray and PCR, 5 in both microarray and RNASeq, none in RNASeq and PCR, and 6 were upregulated in all three platforms (**Fig. 3.3B**). For downregulated DEmiRs, 19 were downregulated in both microarray and PCR, 2 in both microarray and RNASeq, and 3 in RNASeq and PCR, and 4 downregulated in all three platforms (**Fig. 3.3B**). Overall, the meta-analysis showed that within microarrays and PCR, some DEmiRs have contradictory expression in that they are both upregulated and downregulated in the same study. However, RNASeq showed DEmiR expression changing in one direction only, up, or downregulated. Hierarchical cluster analysis of the DEmiR metadata also revealed conflicting data (**Fig. 3.3B**). Some DEmiRs were upregulated in some studies and downregulated in others. This conflicting data was observed in microarray and PCR based profiling, while RNASeq showed expression exclusively in one direction or the other. Furthermore, there were no DEmiRs identified in all the 25 studies analyzed. We observed that DEmiRs such as miR-222, miR-223 and miR-708 were present in at least 36% (9/25) of the studies examined (**Fig. 3.3C** and **Supp file 3.3**). However, these miRNAs showed expression changes in both directions. Others which showed expression in both directions but were present in a lower number of studies include let-7e (16%), miR-34a (24%), miR-100 (20%), miR-27a (16%) and miR-150 (12%) (**Fig. 3.3C**). Overall, there was no clear miRNA consensus of DEmiRs from this meta-analysis.

We next looked at DEmiRs identified in at least 15% (>4/25) of all the studies (**Fig. 3.3D**). Based on this cut-off, we identified DEmiRs that showed consistent expression change in one direction

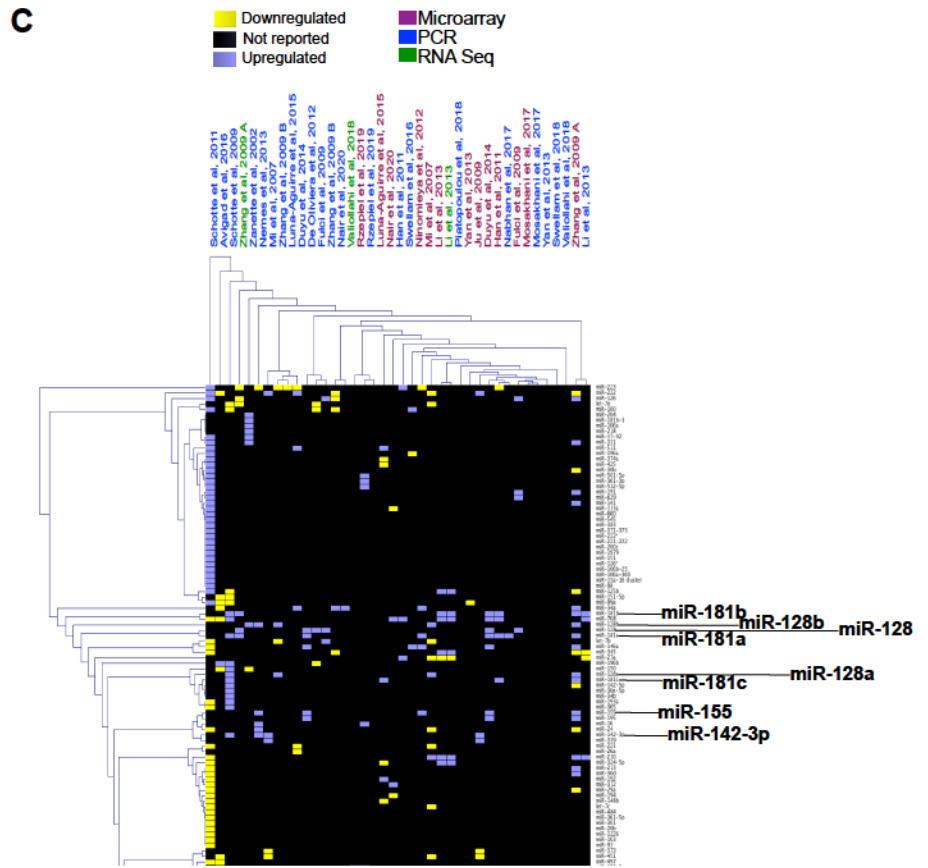
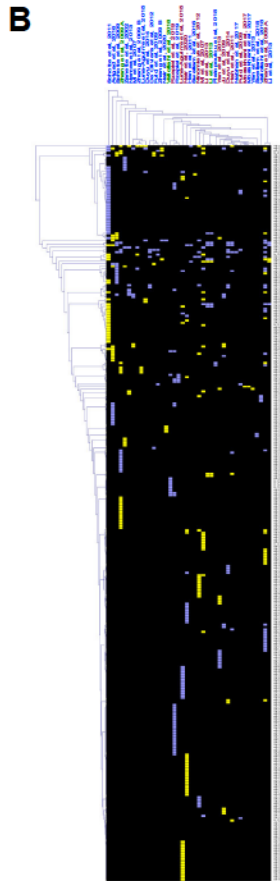
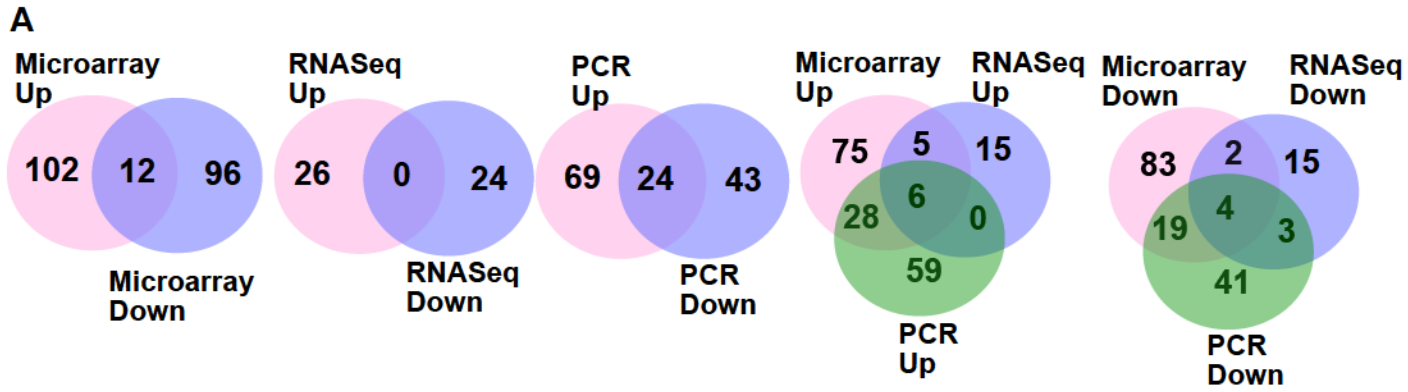


Fig. 3.3. Meta-analysis of differentially expressed miRNAs (DEmiRs) from published papers show a lack of consistency of DEmiRs across multiple platforms.

(A) 2-way Venn diagrams comparing upregulated to downregulated DEmiRs identified from microarray, RNASeq and PCR analyses, respectively. (B) 3-way Venn diagrams comparing upregulated DEmiRs identified from microarray, RNASeq and PCR techniques. (C) Hierarchical cluster analysis of all DEmiRs were organized based on study and profiling technique. Horizontal yellow bars represent downregulated DEmiRs, lilac bars represent upregulated DEmiRs, and black rows indicate DEmiRs not reported. (D) A subset of the hierarchical cluster (from panel C) is shown, with promising DEmiRs indicated.

(**Fig. 3.3D**). We have called these promising DEmiRs. These include miR-181b (32%), miR-128b (20%), miR-181a (32%), miR-128 (24%), miR-128a (16%), miR-181c (16%), miR-155 (20%), miR-142-3p (16%) and miR-451 (16%). These are also listed in **Table 3.1**.

3.2.3 Target prediction identifies mRNA targets.

To identify potential mRNA targets of promising DEmiRs, the 8 DEmiRs were analyzed using miRwalk and miRDB databases. We identified the mRNA targets that were identified by both miRwalk and miRDB to reduce the probability of false positives. There were between 1 and 191 overlapping mRNA targets identified for these 8 DEmiRs (**Table 3.1**). Of the overlapping mRNA targets, the top ranked 25 (based on miRwalk ranking) were selected for pathway and network analyses (**Supp file 3.4**). Biological processes and pathway networks were generated for each miRNA target set (**Fig. 3.4**), summarized in **Fig. 3.5**. Overall, the top putative targets of the 8 promising miRNAs identified in this study are associated with a range of biological processes and pathways including B cell signaling, stem cell differentiation, neuronal cell development and activity, cell cytoskeleton and remodeling cell metabolic activity, gene regulation, innate and adaptive immune cell activity. The notable biological and gene pathways of miR-128-1-5p (**Supp. file 3.5**) targets included negative regulator of I- κ B phosphorylation, voltage-gated chloride channel, neurotransmitter receptor localization, methylenetetrahydrofolate reductase activity, D-serine biosynthetic processes, nephrin family interactions, enzymatic degradation of dopamine by COMT, mannose type O-glycan biosynthesis and RUNX2 transcription factor activity. mRNA targets of miR-128-2-5p (**Supp. file 3.6**) were found to be involved in T and B cell primary

Table 3.1. Top promising genes from metanalysis and re-analysis. Columns show the list of miRNAs, their expression pattern, the name of the mature transcript, the number of predicted targets (miRbase and miRwalk prediction tools) and the overlapping targets.

	miRNA	Expression	Mature transcripts (miRbase)	# of targets (miRwalk)	# of targets (miRDB)	# of overlapping targets (miRwalk vs miRDB)
1	hsa-miR-128	Upregulated	defunct	-	-	-
2	hsa-miR-128a	Upregulated	hsa-miR-128-1-5p	2175	189	37
3	hsa-miR-128b	Upregulated	hsa-miR-128-2-5p	1984	180	42
4	hsa-miR-142-3p	Upregulated	hsa-miR-142-3p	784	739	59
5	hsa-miR-155	Upregulated	hsa-miR-155-5p	483	701	44
6	hsa-miR-181a	Upregulated	hsa-miR-181a-5p	621	1408	118

7	hsa-miR-181b	Upregulated	hsa-miR-181b-5p	1307	1408	191
8	hsa-miR-181c	Upregulated	hsa-miR-181c-5p	556	1409	92
9	hsa-miR-451	Downregulated	hsa-miR-451a	344	40	1

Fig. 3.4. Network analyses of biological processes and gene pathways.

Network analyses for mRNA targets for 8 promising miRNAs that represent both biological process and genetic pathway network. For each miRNA, triangles indicate pathways and circles indicate biological processes while the colored label for each node indicates the overall process.

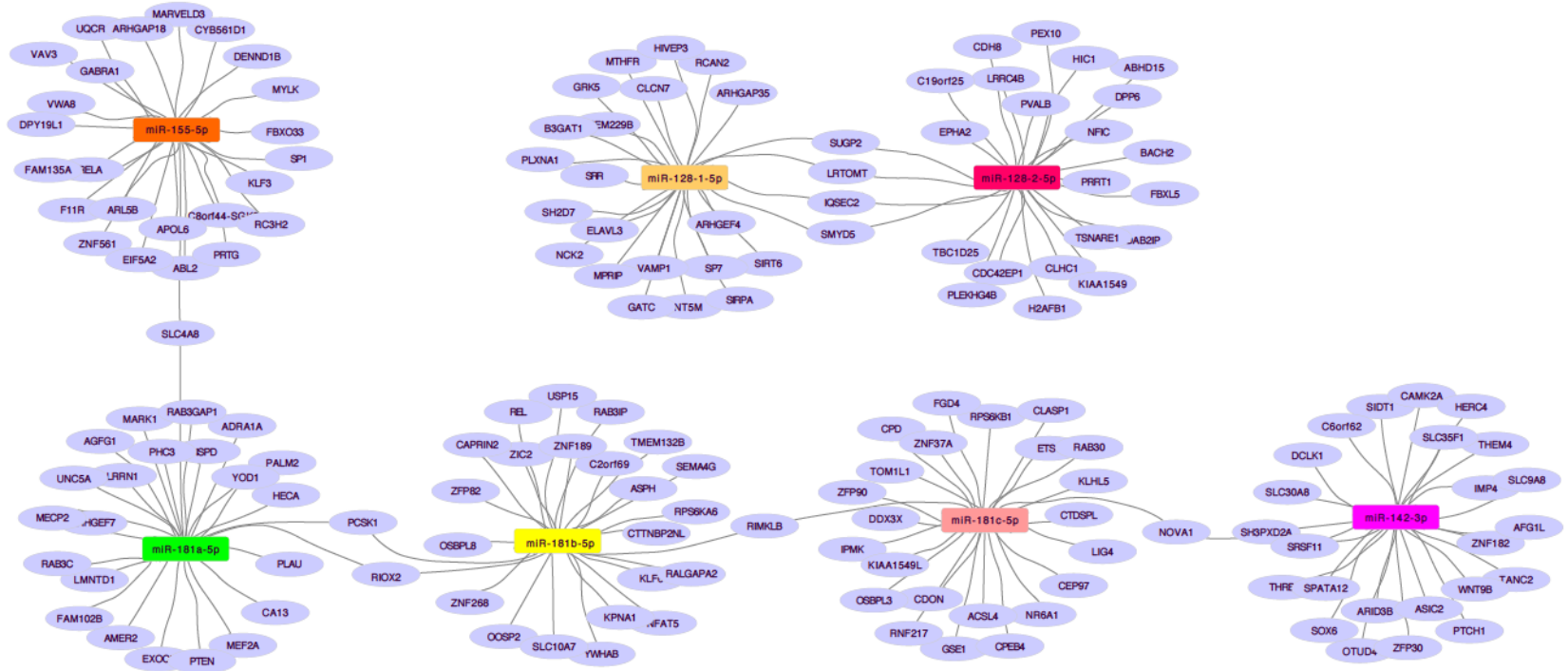


Fig. 3.5. Interaction between miRNA predicted targets. The predicted targets for miR-155-5p and miR-181a-5p interact, while miR-181a-5p and miR-181b-5p targets interact. miR-181b-5p and miR-181c-5p targets interact, while miR-181c-5p targets interact with miR-142-3p. Finally, miR-128-1-5p and miR-128-2-5p targets interact.

adaptive immune responses, ARF protein signal transduction, regulation of toll-like receptor 4 signaling, as well as pseudopodium organization. Other miR-128-2-5p targets include postsynaptic specialization assembly, enzymatic degradation of dopamine by COMT, termination of RNA Polymerase III transcription, notochord cell differentiation, pseudopodium organization and transcription factor SUMOylation. miR-155-5p (**Supp. file 3.6**) targets include regulators of miRNA metabolic process, T cell cytokine production, actin cytoskeleton reorganization, Rac guanyl-nucleotide exchange factor activity, smooth muscle development, and T-helper 2 cell cytokine production. Other targets of miR-155-5p are involved in protein C-linked glycosylation, positive regulation of neurotransmitter secretion, polyamine homeostasis and ubiquinol-cytochrome-c reductase activity. miR-142-3p (**Supp file 3.7**) targets were regulators of transporter activity including metal ion protein antiport, nucleic acids transmembrane transport and zinc and metal ion transmembrane transport. miR-142-3p targets also included regulators of collecting duct development, female courtship behavior, negative regulation of cardiomyocyte differentiation, Cytochrome c electron transport, calcium dependent protein kinase activity and CD28-dependent PI3K/Akt signaling. miR-181a-5p (**Supp. file 3.9**) targets include those involved in nervous system processes, regulation of microtubule cytoskeleton, positive regulation of neurotransmitter secretion, and insulin processing. Other miR-181a-5p targets include plasminogen activation, histone demethylase activity, SUMOylation of DNA methylation, mannose type O-glycan biosynthesis, catechol-containing compound biosynthesis, netrin receptor activity and positive regulation of dendritic spine morphogenesis. miR-181b-5p (**Supp. file 3.10**) regulates mTORC1-mediated signaling, phospholipid transfer activity, suppression of phagosomal maturation, canonical NF-KB pathway, and cancer immunotherapy to PD-1 blockade. In addition, peptide-aspartate beta-dioxygenase activity, glycoprotein transport, RSK activation, suppression of

phagosomal maturation, incretin biosynthesis and histone deubiquitination are targets of miR-181b-5p. miR-181c-5p (**Supp. file 3.11**) regulates mTORC1-mediated signalling targets, positively regulates protein autophosphorylation, fatty-acyl-CoA metabolic process, transcriptional regulation of pluripotent stem cells, and inositol-1,4,5 triphosphate 6-kinase activity. Furthermore, embryonic body morphogenesis, response to lithium ion, mitotic spindle assembly regulation, negative regulation of cytoplasmic translation and glutamate/glutamine metabolism are targets of miR-181c-5p. miR-451a (**Supp. file 3.12**) positively regulates processes involved in myoblast formation, positive regulation of transforming growth factor beta production, regulation of cardiac muscle proliferation and activation of AP-1 family of transcription factors.

3.3 Discussion

The objective of this study was to identify a putative miRNA signature of B-ALL, which could be validated as a biomarker for better tracking of disease progression and response to therapy. To do this, we performed a two-pronged systematic analysis of the miRNA landscape in B-ALL: a reanalysis of B-ALL miRNA profile raw data and a meta-analysis of DE miRs identified in the literature. Specifically, for the reanalysis approach, we identified 5 datasets, with expression data of over 500 distinct miRNAs in 200 B-ALL patient and corresponding healthy control samples. Concurrently, we identified 25 published papers with over 100 DE miRs between B-ALL patient and HC samples. Though our results showed no clear miRNA consensus signature in B-ALL, we identified 8 promising miRNAs from the meta-analysis whose predicted mRNA targets may potentially be relevant components of BCR signaling pathways in B cells.

Unsupervised hierarchical clustering of all miRNAs across all platforms showed that miRNAs clustered by platform or dataset, rather than by disease state. This suggests that platforms have a

stronger effect on miRNA profiling than samples. This strong platform effect may be due to RNA isolation method, similar to what was found in a previous study showing that the diversity of RNA isolation kits is potentially a large source of variability between miRNA profiling datasets⁴⁷⁷. In addition, RNA isolation techniques may be biased towards enriching for specific miRNAs⁴⁷⁸. Variability in miRNA detection across platforms may also be linked to profiling platforms lacking sufficient sensitivity to show miRNA concentrations at low levels, in turn skewing statistical differences in expression⁴⁷⁹. Furthermore, different laboratory procedures, design and manufacture of arrays, hardware detection, algorithms for extraction of intensity signals and bioinformatics analyses could further exacerbate these inter-platform miRNA signaling discrepancies⁴⁷⁹. Thus, these analyses clearly demonstrate the need for the development of standardized sample isolation procedures to allow for meaningful comparison and integration of data from independent studies.

In our reanalysis cohort, HC, and B-ALL samples both express miRNAs known to regulate different B cell subsets during development, differentiation, and activation. Of the >700 miRNAs profiled across 5 datasets in HC and B-ALL samples, less than 200 miRNAs were differentially expressed between both states. A putative signature of 12 DE miRs from this approach were miR-130b, miR-181b, miR-181c, miR-181d, miR-34a, miR-143, miR-145, miR-223, miR-134, miR-27a, miR-593, miR-650 (HC versus B-ALL subtypes). However, these DE miRs were not common to all 5 datasets, and/or were expressed in different directions (up in one and down in the other). Therefore, these DE miRs do not form a signature for B-ALL. Another 8 DE miRs were common to 3/5 datasets – miR-22, let-7d, miR-195, miR-223, miR-130b, miR-24, miR-186 and miR-34a. These genes which all have known links to normal hematopoiesis and targeting cancer associated genes in different disease. Therefore, these DE miRs are showing to be promising in B-ALL, but not powered enough to be strongly considered as signatures.

Meta-analysis of DE miRNAs profiled by microarrays, RNASeq, and PCR showed that some miRNAs were upregulated in some studies and downregulated in others, while others were both upregulated and downregulated in the same study. A potential source for this discrepancy could be the highly heterogeneous nature of B-ALL as there are many subtypes of B-ALL, as previously described. One additional layer of B-ALL disease heterogeneity is the age- and sex-associated difference in etiology, clinical presentation, response to therapy⁴⁸⁰. For instance, pediatric B-ALL, which is the most prevalent leukemia in children, differs significantly from adult B-ALL. Pediatric B-ALL patient samples have been shown to have significant heterogeneity in morphology, immunophenotype, genetic aberrations and therapeutic response⁴⁸¹. Furthermore, B-ALL contains multiple coexisting B-ALL subclones, which are immunophenotypically heterogeneous leukemia populations with bimodal or broad expression of surface markers⁴⁸¹. Accordingly, miRNA expression levels may vary between subtypes, and age groups, which makes profiling of miRNA in B-ALL without proper subtyping give rise to conflicting results^{482,483}. For instance, in our reanalysis cohort, there was more discrepancy in our non-subtyped B-ALL datasets (GSE58199, GSE2564 and GSE56584) than between the subtyped datasets (GSE31376 and GSE23024). Of the meta-analysis cohort, 16% (4/25) papers were non-subtyped, which showed both down- and up-regulation of the same genes. For the subtyped datasets, some miRNAs may be changed in one direction in one subtype and in another in a different subtype. Taken together, miRNA profiling of B-ALL should be approached in a subtype specific manner, at the very least to improve the ability to effectively characterize and compare in a manner that acknowledges the subtype heterogeneity.

A key consideration for immunophenotyping B-ALL miRNAs is the type of controls used. The correct controls for profiling B-ALL miRNAs should be CD19 enriched cells (B cells), rather than

total PBMCs. This is because in peripheral blood and bone marrow, B cells constitute a small percentage of the total PBMCs (5-10%)⁴⁸⁴. Therefore, if the controls are not B cell enriched, this could skew the balance of DE miRs between datasets. This is another limitation of the studies profiled here.

B-ALL originates from normal B cells, which undergo transformation into lymphoblasts that are phenocopies of different stages of B-cell development⁴⁸⁵. Consequently, the identified promising DE miRs from the meta-analysis regulate specific B cell development relevant genes. For instance, miR-142 and miR-128-2 regulate B cell development at the pre pro-, pre-, pro-, immature and circulating stages with miR-128-2 targeting MALT1, ERK and MAPK associated signaling^{486,487}. Some of the promising DE miRs are members of the miR-181 family (miR-181-a1/b1, miR-181-a2/b2 and miR-181-c/d clusters)⁴⁸⁸. The miR-181 family are expressed in hematopoietic cells, with dynamic changes to their expression; being high in the early B-cell stage, and declining progressively⁴⁸⁹. Furthermore, the miR-181 family and miR-155 regulates germinal centre B-cell differentiation, via targeting of activation-induced cytidine deaminase and promotes generation of chromosomal c-myc/IgH translocations, which are important in leukemia onset including in a rare B-ALL subtype called Burkitt-type ALL⁶¹. In addition, three of 12 DE miRs from the reanalysis cohort (miR-181b and miR-181c) were also part of the promising DE miRs. While this may strengthen the case for their potential as putative miRNA signatures of B-ALL, their expression levels were contradictory in the GSE31376 versus GSE23024 dataset (**Fig. 3D**).

miRNA target identification is a first step to elucidating miRNA functions in different relevant signalling pathways⁴⁹⁰. The *in-silico* target prediction identified hundreds of putative targets, which showed distinct patterns when further narrowed down. Some of these patterns have previously been shown in B-ALL. For instance, NF- κ B, a transcription factor regulated by I- κ B,

a target of miR-128-1-5p, is constitutively activated in pediatric B-ALL, irrespective of subtype and facilitates abnormal B cell proliferation⁴⁹¹. Furthermore, loss of I- κ B phosphorylation regulation, identified as a target network of miR-128-1-5p in our study, is relevant as a hallmark of cancer⁴⁹². NF- κ B signalling, also a target of miR-181b-5p, is upregulated in pediatric B-ALL, which alters the action of CCCTC-binding factor (CTCF), a candidate tumour suppressor⁴⁹³. Furthermore, toll-like receptor 4 signalling, a target of miR-128-2-5p, was identified as necessary in driving B-ALL immune responses in pediatric B-ALL cell lines⁴⁹⁴. Another target network of miR-128-2-5p is transcription factor SUMOylation. SUMOylation of Ikaros, a transcription factor and tumour suppressor known to be important in lymphocyte development, was shown to be abundant in B-ALL but absent in healthy peripheral blood leukocytes⁴⁹⁵.

Cell transporter networks that were found to be targeted by miR-142-3p may play a role in B-ALL pathogenesis. For instance, zinc ion homeostasis, which is essential for immune cell function, is regulated by transporters, which, when dysregulated facilitates B-ALL pathogenesis and acts as a mediator of several pathways, including immune cell homeostasis⁴⁹⁶. Also, protein kinase C (PKC), a superfamily of serine/threonine kinases, another target of miR-142-3p, is implicated in the alteration of hematopoietic stem cell (HSC) maturation⁴⁹⁷. Specifically, the overexpression of PKC- α suppressed the expression of mitochondrial protein phosphatase A (PPA) activity in the B-ALL cell line REH by partly targeting the expression of Bcl2 and subsequently mediating chemoresistance⁴⁹⁸. TGF-beta signalling, a target of miR-155-5p, inhibits the growth of pre-B-ALL cell line BLIN, which identified TGF-B as a potential anti-growth cytokine for B-ALL⁴⁹⁹. Phosphatidylinositol 3-kinase/Protein Kinase B (PI3K/Akt signalling), another target of miR-155-5p, is linked to stromal cell-mediated survival of ALL cells⁵⁰⁰. miR-181c-5p targets inositol-1,4,5

triphosphate-6 kinase which phosphorylates and regulates IP₃ to IP₄ and plays a role in Ca²⁺ homeostasis/signalling which is essential for normal B cell responses⁵⁰¹.

A target of miR-181a-5p is ERK/MAPK signalling, with negative regulation of ERK correlating with pre-B cell transformation, while pathway activation due to point mutations is also a hallmark of B-ALL⁵⁰². mTORC1 signalling, a target of miR-181b-5p and miR-181c-5p, is a signalling complex associated with the Mammalian/mechanistic target of rapamycin (mTOR), a conserved serine/threonine kinase that is deregulated in B-ALL, thus making it a therapeutic candidate for B-ALL⁵⁰³. Finally, the AP-1 family of transcription factors, which are targets of miR-451a, are identified as a crucial factor in leukemia development and associated with glucocorticoid resistance⁵⁰⁴.

The network analyses showed an enrichment of innate immune and B cell receptor associated signaling pathways. This suggests that the promising DE miRNAs in B-ALL target and dysregulate signaling and biological pathways relevant for normal B cell function, which solidifies the role of miRNAs as drivers of early hematopoietic differentiation. This further suggests that these DE miRNAs may be suitable candidates for further research on miRNA in B-ALL monitoring, diagnosis, or progression. Additionally, there is also a preponderance of neuronal cell, stem cell, transcription regulation, fatty acid metabolism and cell cytoskeleton remodeling relevant biological processes and gene pathways. These additional processes and biological pathways have not previously been shown or associated with these miRNAs, and questions remain as to the link between these miRNAs and their targets in B-ALL. Overall, the predicted interactions between these 8 promising miRNAs cannot be used to draw actionable conclusions and need to be investigated further to fully determine the actual roles of these miRNAs in B-ALL onset and progression.

3.4 Conclusion

To reconcile the myriad data profiling of miRNA in B-ALL and identify an actionable miRNA signature, this study sought to define a signature of DEmiRs from reanalysis of microarray data and meta-analysis of the published literature. Overall, despite the abundance of information on DEmiRs, there is no clear consensus signature DEmiRs for B-ALL. However, we found 8 genes that are promising DEmiRs common to all 5 reanalysis datasets. These promising genes have known pro-cancer and pro-leukemia links which make them significant. Thus, to move the field forward, a combination of biological and technical variables need to be standardized, to ensure that DEmiRs can be profiled with less lab-to-lab and inter-platform/technique variations.

Chapter 4: Extracellular vesicle small RNA cargo discriminates non-cancer donors from pediatric B cell acute lymphoblastic leukemia patients.

4.1 Introduction

B-cell acute lymphoblastic leukemia (B-ALL) is the most prevalent pediatric malignancy in Canada and worldwide, which, together with T-cell acute lymphoblastic leukemia, make up 75 – 80% of all new pediatric cancers, contrasting with 20% of all adult leukemias^{240,505}. In B-ALL, developing B cells at different stages encounter development arrest, leading to abnormal growth and survival¹⁷². Clonal proliferation of growth-arrested B cells, called lymphoblasts, in the bone marrow (BM) is soon followed by extramedullary migration and lymphoblasts circulating in the blood, which crowd out healthy cells^{159,173}. Overall, chromosomal changes and molecular changes in B cell gene expression and regulation leads to changes in B cell development processes, replace normal B lymphocytes with lymphoblasts in the BM and extramedullary sites. Due to clonal expansion, the abundance of lymphoblasts compromises the lymphocyte population at the immune sites, leading to impaired immune function.

In B-ALL, specific genetic aberrations at different stages of B cell development facilitate risk stratification and prognosis. There are two groups of common B-ALL-associated genetic abnormalities. The first is numerical abnormalities (Hyperdiploidy and hypodiploidy) with loss or gain of chromosomes. In the second group, structural abnormalities (such as *E2A/TCF::PBX1*, myeloid/lymphoid or mixed-lineage leukemia [MLL] rearrangements, also known as KMT2A, *BCR::ABL1* [Ph-positive], Ph-like and *ETV6::RUNX1*) contain translocations that encode chimeric proteins or move a gene close to a strong transcriptional promoter resulting in gene overexpression^{160–162}. Of these cytogenetic subtypes, the most favourable (low-risk) and most

prevalent subtypes for pediatric ALL are ETV6::RUNX1 and Hyperdiploidy, while high-risk subtypes such as Ph-like and KMT2A have the worst outcome and a low prevalence in pediatric patients ⁵⁰⁶. In addition to cytogenetic subtype-based risk stratification, other prognostic factors include patient age, central nervous system (CNS) involvement and response to induction/consolidation therapy ¹⁵⁴.

The cure rates and survival outcomes, especially for pediatric ALL, have improved tremendously, largely due to molecular genetics, improved testing for measurable residual disease (MRD), risk-based treatment regimens, new targeted agents and allogeneic hematopoietic stem cell transplantation (HCT) ^{154,507}. This has led to 5-year overall survival (OS) rates of 56% for infants, 89% for children, 61% for adolescents and young adults, and 20 – 40% for adult patients ⁵⁰⁷⁻⁵¹¹. Despite the intense chemotherapy regimens, including CNS prophylaxis, approximately 20% of B-ALL patients suffer a relapse ⁵¹². Currently, the strongest independent prognostic factor of pediatric B-ALL relapse is MRD, which is low-level disease that is indicative of disease burden and is undetectable by conventional cytomorphology ⁵¹³. MRD detection post-initial treatment correlates with poorer relapse-free survival (RFS) and OS ^{514,515}. Specifically, MRD positivity during standard chemotherapy, especially after the end of consolidation, is associated with an increased risk of relapse; thus, early detection would enable salvage treatment to be initiated earlier, which could possibly improve treatment results ^{513,516}. Conventional methods for detecting MRD involve invasive bone marrow aspirates followed by polymerase chain reaction (PCR) and flow cytometry-based detection of abnormal lymphoblasts. However, less invasive methods to obtain samples to track therapy response and MRD could improve the frequency of monitoring of disease burden and enable real-time clinical decision-making.

Extracellular vesicles (EVs) are membrane-enclosed nanoparticles released by all living cells tested to date ^{246,517}. Based on the size and mode of biogenesis, EVs are classified into three main groups, which significantly overlap in size and composition. Exosomes are 30-150nm-sized EVs produced via the endocytic pathway ⁵¹⁸. Microvesicles are 100-1000 nm-sized vesicles that are formed via the outward budding of the plasma membrane (PM) ^{517,519}. Apoptotic bodies are 1000-5000 nm-sized particles produced during the final stages of apoptosis. During EV biogenesis, cell-derived biomolecules are shuttled into the EVs, resulting in bioactive cargo (proteins, lipids, metabolites and nucleic acids), with cargo content being representative of the originating cell ²⁴⁶. Thus, a unique and important characteristic of EVs is their cargo, which affects other aspects of EV biology, especially their function. For instance, EV surface protein cargo can act as ligands to interact with receptors on other cells and facilitate their uptake and downstream effects ^{246,276}. In addition, EV miRNA content, especially when EVs are taken up by other cells, can stimulate other functions in the recipient ⁵²⁰. Hence, examining the cargo profile of EVs can be informative about the physiological or pathological state of the originating cell or the function of the EVs.

Nucleic acids that can be packaged into EVs include mRNA and microRNA (miRNA)^{293,521}, long non-coding RNA (lncRNA) ⁵²², as well as ribosomal RNA (rRNA)^{293,523}. EVs have also been reported to carry genomic DNA (gDNA) and mitochondrial DNA (mtDNA); however, this has recently been challenged ^{412,524,525}. The nucleic acid profile of cancer patient EVs has been used to identify signatures that distinguish them from non-cancer donors (NCD). Specifically, EV RNA signatures have been identified for cancers including pancreatic cancer ⁵²⁶, ovarian cancer ⁵²⁷, clear cell renal carcinoma ⁵²⁸ and breast cancer ⁵²⁹, acute myeloid leukemia ⁵³⁰, and chronic lymphocytic leukemia ^{365,421}. To the best of our knowledge, the small RNA EV content of pediatric B-ALL has not been reported.

EVs hold the promise for filling in the gaps about pediatric B-ALL aetiology and for less invasive monitoring of disease progression. Specifically, B-ALL EV physical characteristics (size and concentration) and biological properties, especially RNA content, have shown promise in other diseases, including leukemia. In pediatric B-ALL patients, the most relevant biofluids that carry clinically actionable biomarkers are peripheral blood (PB) plasma, bone marrow aspirate and cerebrospinal fluid. In pediatric B-ALL PB plasma, the size and concentration of EVs have not previously been measured, especially in comparison to age-appropriate NCD. Furthermore, RNA transcript types such as miRNA and mRNAs, which have been previously linked to pediatric B-ALL, have not been clearly explored in pediatric B-ALL EVs. Therefore, we have explored pediatric B-ALL EVs (n = 6) at diagnosis, in comparison to NCD EVs (n = 6), as a proof-of-principal study to determine if EVs could be a source for easily accessible biomarkers.

4.2 Results

4.2.1 Pediatric B-ALL plasma contains more extracellular vesicles than NCD.

Nanoparticle tracking analysis (NTA) was used to profile extracellular vesicles (EVs) in pediatric B-ALL and NCD peripheral blood plasma. EV diameter for both B-ALL and NCD is 75 – 120 nm (average) (**Fig 4.1A, B and Fig S1**). A comparison of the mean and mode diameters (pediatric B-ALL versus NCD) showed that there is no statistically significant difference between NCD and pediatric B-ALL EV sizes (**Fig. 4.1A, B**). A comparison of the EV concentration (number of particles/ml) of the NCD samples showed that NCD EV concentrations are almost uniform (**Fig 4.1C**). In contrast, B-ALL EV concentrations showed high sample-to-sample variability, suggesting that the pediatric B-ALL samples are highly heterogenous (**Fig 4.1C**). Pediatric B-ALL peripheral blood plasma has a higher EV concentration than NCD (**Fig. 4.1C**).

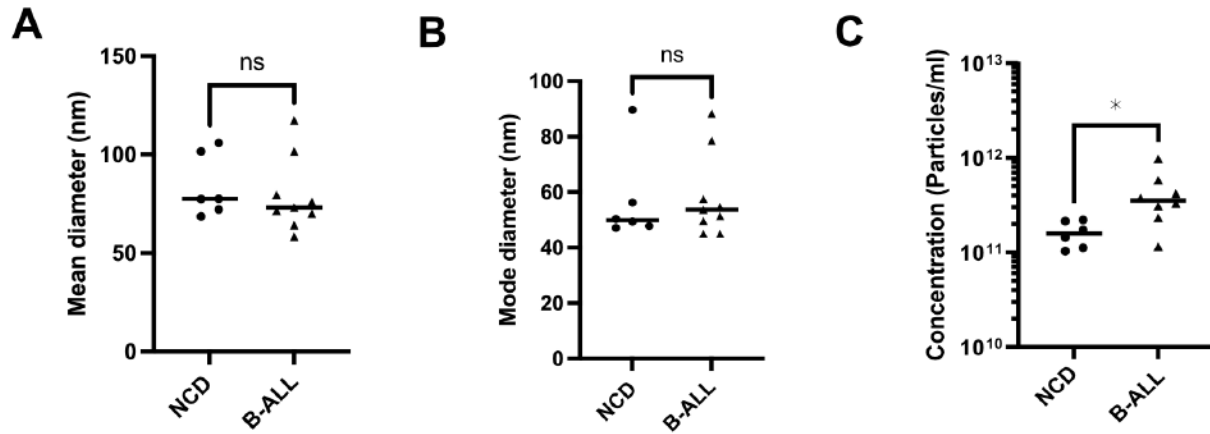


Fig. 4.1. Characterization of EVs in non-cancer donors (NCD) and pediatric B-ALL blood plasma using nanoparticle tracking analysis showed a higher EV concentration in B-ALL. (A) Mean diameter, (B) mode diameter, and (C) concentration of plasma from NCD and pediatric B-ALL plasma. Statistical significance determined by t-test (* $p < 0.05$).

4.2.2 Vn96 isolates a more comprehensive EV population than SEC or ExoQuick

In the absence of a gold standard EV isolation technique that is clinically amenable for EV analysis, we first evaluated commercially available EV isolation techniques to identify the best method for our study. For our purposes, the criteria for clinical amenability include having an easy protocol with applicability in the clinic, compatibility with small volumes (typically <5 ml), good EV yield, low co-isolation of contaminating proteins, and not requiring any specialized equipment. Using available literature and previous experience in our lab, we identified size exclusion chromatography (SEC) and polyethylene glycol (PEG) polymer-based ExoQuick as increasingly popular isolation techniques suitable for our purposes⁵³¹. We also identified the synthetic peptide Vn96 as appropriate for our purposes due to the ease of use and previous experience^{270,351}.

First, we isolated EVs from the conditioned culture media (CCM) of pediatric B-ALL cell lines RCH-ACV and UoC-B1, followed by Western blot analysis to characterize the EVs. We probed for EV-specific markers - tetraspanins CD63 and CD81 and heat shock proteins HSC70 and HSP90; B cell marker BLNK; and the contamination marker Calnexin (endoplasmic reticulum marker). The results showed that all three isolation methods isolate EVs from CCM but with different cargo (**Supp. fig. 4.2A-C**). Specifically, ExoQuick (**Supp. fig. 4.2A**) and SEC (**Suppl. fig. 4.2B**) isolate EV subpopulations from pediatric B-ALL CCM that are enriched for CD63 or CD81, respectively. In contrast, Vn96 isolates a more comprehensive EV population, as shown by the inclusion of multiple EV markers, including CD63, CD81, HSP70 and HSP90 (**Suppl. fig. 4.2C**). EVs were also isolated from healthy donor plasma and probed for the above-described EV markers, as well as platelet activation marker CD41 and markers of contamination for plasma: Albumin, Apolipoproteins A (High-density lipoprotein), Apolipoprotein B (low-density lipoprotein) and CD235a (Erythrocytes). EVs isolated from plasma using Vn96 showed a similar

protein marker profile as the Vn96-isolated EVs from CCM, with a low signal for plasma contamination markers (**Supp. fig.4.2D**). Finally, the Vn96 isolated EVs have a round morphology, showing a characteristic double membrane ⁴¹⁷ (**Supp. fig. 4.3**). Therefore, we selected Vn96 for subsequent EV isolations.

4.2.3 NCD and pediatric B-ALL EVs package different types of small RNA.

We next performed small RNA sequencing analysis (small RNA-seq) on RNA extracted from EVs that were isolated from B-ALL plasma compared to NCD plasma. sRNADE-based annotation of the overall RNA transcripts present in NCD and pediatric B-ALL EVs revealed that many different types of RNA transcripts are packaged into EVs with 11.3% +/- 7.8% unassigned (**Fig. 4.2A**; unassigned not shown). Packaged RNA transcripts include miRNAs, lncRNAs, tRNAs, mRNAs, and snRNAs. The most abundant types of RNA transcripts are miRNAs and fragments of yRNAs; however, miRNA made up a minority of the total RNA abundance in most cases and only becomes a majority after excluding the unassigned RNA transcripts. We found that variable miRNA transcript distribution between NCD and pediatric B-ALL exists (**Fig 4.2A**). The difference in abundance is statistically significant (unpaired t-test, $p < 0.05$) for lncRNA, sense mRNA, other RNA, rRNA, snoRNA, snRNA, tRNA and yRNA, with only yRNA decreased in B-ALL and all other RNA species showing an increase. Closer inspection of the RNA transcripts suggests that many are fragments, as is consistent with the isolation of small RNA, which was verified by the analysis on the Agilent fragment analyzer. Some earlier publications reported that miRNAs are the most abundant RNA transcripts in EVs ^{293,532}; however, we and others ^{526,533} have not reproduced this finding when unassigned transcripts are considered.

4.2.4 A miRNA-specific signature does not clearly distinguish NCD from pediatric B-ALL

The differentially packaged miRNAs (DPmiRs) between NCD and B-ALL EVs were analyzed using the miRNA-specific sRNADE pipeline. DPmiRs (FDR<0.05) that were with all five platforms were analyzed by unsupervised hierarchical cluster analysis (**Fig 4.2B**). miRNAs downregulated in B-ALL EVs include miR-758-3p, miR-335-5p, miR-26b-5p, miR-340-3p, and let-7f-5p, while miRNAs such as miR-455-3p, miR-128-3p, miR-181a-3p, miR-181b-5p and miR-1246 are upregulated in most B-ALL samples and downregulated in 3 of 8 B-ALL patient samples. Both mature and precursor miRNAs were detected in all samples (**Table 4.1**); however, analysis of these transcripts did not clearly distinguish NCD from B-ALL samples (**Fig 4.2B**).

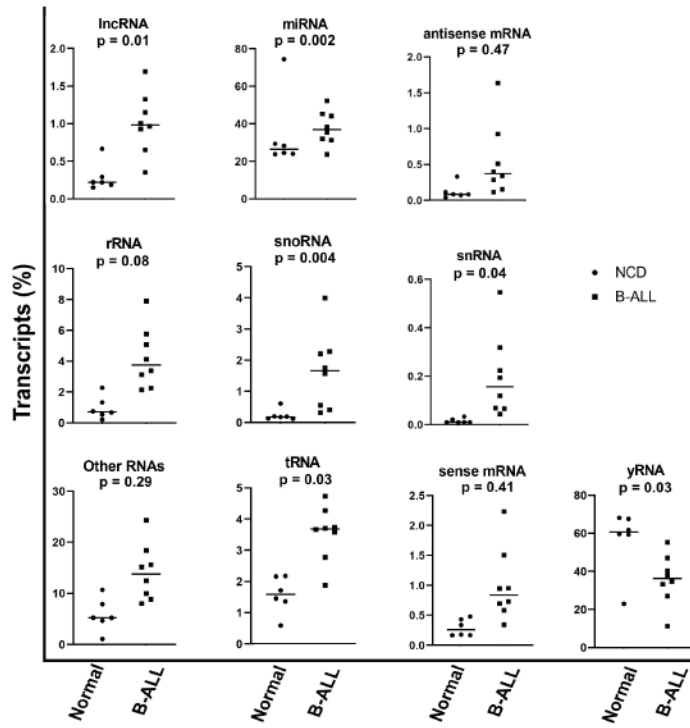
4.2.5 Sample heterogeneity of plasma EVs

As there was a heterogenous mixture of small RNAs detected, we next analyzed the small RNA-seq data in an agnostic manner (i.e., analysis of all transcript types). MA plots, a scatterplot of average expression signal (A -x axis) and log₂ fold change (M – y axis) showed a higher distribution of points below 0 on the y-axis, suggesting a higher number of genes with decreased packaging than increased packaging (**Fig 4.3A**). The principal component analysis showed that most NCD (control) EVs clustered closely together (**Fig 4.3B**). In contrast, B-ALL (disease) EVs tended to have greater variance on the PCA2 axis, demonstrating clear heterogeneity between patients.

4.2.6 EV RNA transcripts discriminate pediatric B-ALL from NCD

After analyzing all RNA transcripts detected in NCD and pediatric B-ALL EVs, we compared all

A



B

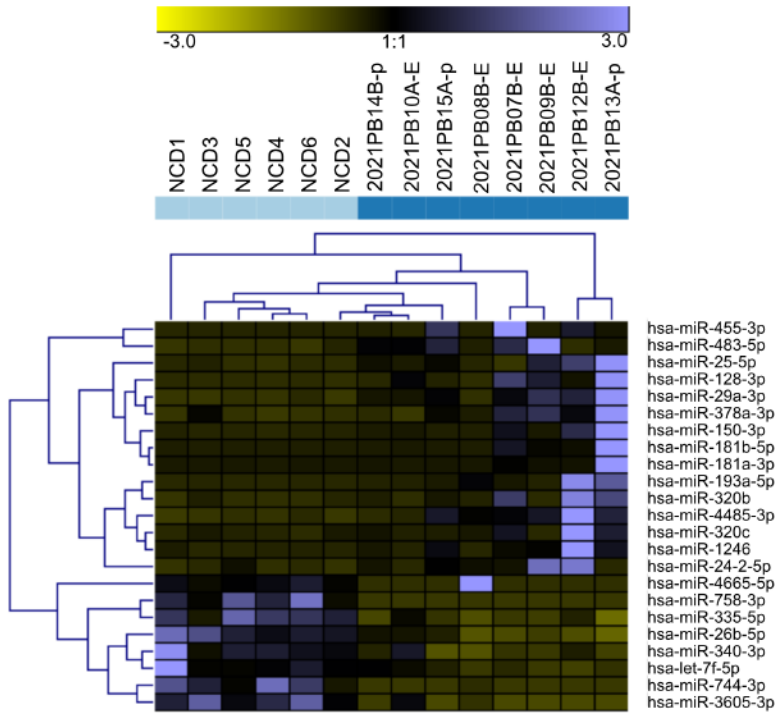


Fig. 4.2. A miRNA-specific EV signature does not distinguish pediatric B-ALL from non-cancer donors (NCD). Raw small RNASeq data was analyzed using the sRNADE platform (SRNA toolbox). (A) Eleven RNA transcript species groups packaged into NCD, and pediatric B-ALL plasma EVs were identified. (B) Unsupervised hierarchical cluster analysis using the sRNADE platform and identified differentially packaged miRNA signatures (FDR <0.05).

transcripts in NCD to those detected in B-ALL. We found that 7738 transcripts were present in both states (**Fig. 4.4A**). We explored these transcripts further to determine if they are differentially packaged in B-ALL EVs compared to NCD EVs. Using DESeq2, we identified differentially packaged small RNA (DPRNA) that includes mRNAs, lncRNAs, tRNA, snRNAs, and snoRNAs (**Fig. 4.4B**). A gene signature was clear when the top DPRNAs (FDR <0.0001) were viewed using unsupervised hierarchical clustering analysis (**Fig. 4.4B**). The clustered transcripts include lncRNAs and mRNAs with only one miRNA transcript (miR-4645) making the cut-off. ETV6::RUNX1 and Ph-like samples were found in both B-ALL clusters.

We previously identified 8 promising genes that were differentially expressed in B-ALL compared to healthy cells, with the premise that these genes could be explored in EVs. These genes – miR-128-1, miR-128-2, miR-142-3p, miR-155, miR-181a/b/c and miR-451a. These miRNAs all have known links to hematopoiesis at physiological conditions, as well as leukemia including B-ALL. These miRNAs were detected in normal and pediatric B-ALL plasma EVs. However, based on our analyses, we did not find them as part of small EV signatures that discriminate pediatric B-ALL from NCD.

Next, we manually annotated and divided the DPRNA signature (FDR<0.05) into different types of RNA transcripts, followed by unsupervised hierarchical cluster analysis. The mRNA fragments-based signatures included RNA from intronic regions such as NEPRO, EHD2, and RFX1 and from exons such as WDR31, NBPF10, NBPF14, NBPF19, NBPF20 and HLA-B, as well as the miRNA miR-4645. For the cluster of differentially packaged mRNAs (DPmRNAs), NCD and B-ALL formed two clear independent clusters (**Fig. 4.4C**). All NCD samples except NCD6 were very

Table 4.1. Total detected miRNA transcripts

Sample ID	Sex	Age at collection (years)	Donor diagnosis	Source	PB WBC Count (x10 ⁹ /L)	Plasma volume (μl)	EV small RNA (ng)	Read count
NCD1	Male	2	Spinal muscular atrophy	BCCH Biobank	NA ^a	1000	76.64	1.25E+07
NCD2	Male	9	Acute tonsillitis	BCCH Biobank	NA	1000	72.84	1.11E+07
NCD3	Male	3	Spinal muscular atrophy	BCCH Biobank	NA	1000	29.89	2.03E+07
NCD4	Male	17	Epilepsy	BCCH Biobank	NA	1000	18.89	2.35E+07
NCD5	Female	17	Epilepsy	BCCH Biobank	NA	1000	1.73	1.54E+07

NCD6	Female	12	Spinal muscular atrophy	BCCH Biobank	NA	1000	0.05	2.94E+07
2021PB07B_E	Male	11	ETV6::RUNX1 pediatric B- ALL	Janeway	NA	1000	110.15	9.64E+06
2021PB08B_E	Male	9	ETV6::RUNX1 pediatric B- ALL	BCCH Biobank	1.20	480	227.06	1.39E+07
2021PB09B_E	Male	3	ETV6::RUNX1 pediatric B- ALL	BCCH Biobank	24.40	480	314.02	8.44E+06
2021PB10A_E	Female	3	ETV6::RUNX1 pediatric B- ALL	BCCH Biobank	16.20	370	9.28	9.76E+06
2021PB12B_E	Male	2	ETV6::RUNX1 pediatric B- ALL	BCCH Biobank	2.20	460	204.73	7.11E+06
2021PB13A_p	Female	12	BCR::ABL1 like pediatric B-ALL	BCCH Biobank	507	290	457.22	1.93E+07

2021PB14B_p	Male	17	BCR::ABL1 like pediatric B-ALL	BCCH Biobank	3.10	440	38.25	8.15E+06
2021PB15A_p	Female	17	BCR::ABL1 like pediatric B-ALL	BCCH Biobank	49.90	320	1.56	1.16E+07

^aNA- not available

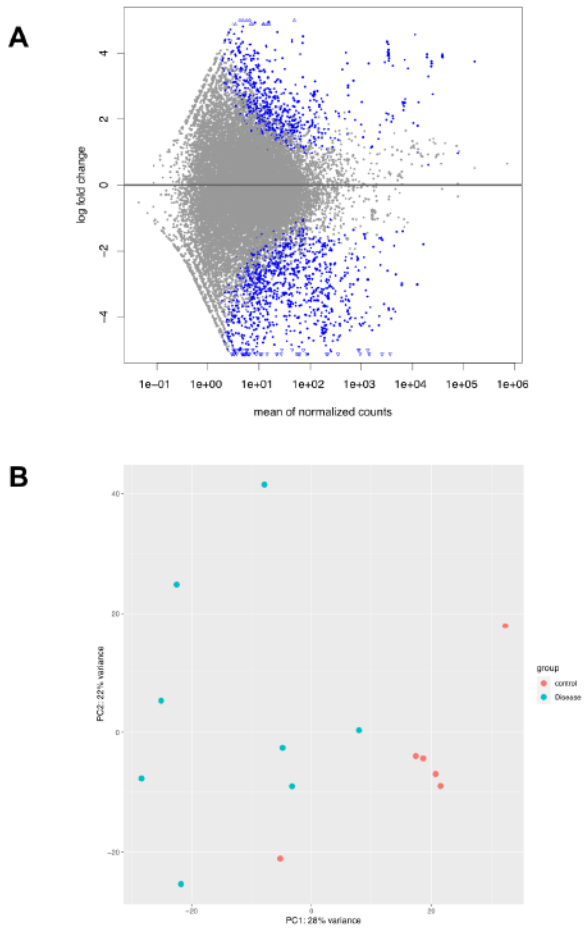


Fig. 4.3. RNASeq data exploration. (A) PCA plots showing clusters of samples based on their similarities. Principal component 1 (PC1) is shown on the x-axis and PC2 on the y-axis. The legend is shown on the right – pink circle – NCD and green circle – B-ALL samples. (B) Volcano plot created using the DESeq2 platform in R shows overall pattern of differential packaging. Black represent RNA transcripts that are not significantly different, blue circles are transcripts that are significantly differentially packaged (FDR < 0.05) and red circles are transcripts significantly differentially packaged (FDR < 0.05) and absolute log fold-change > 2. (C) Unsupervised hierarchical cluster analysis of DPRNA (FDR < 0.0001).

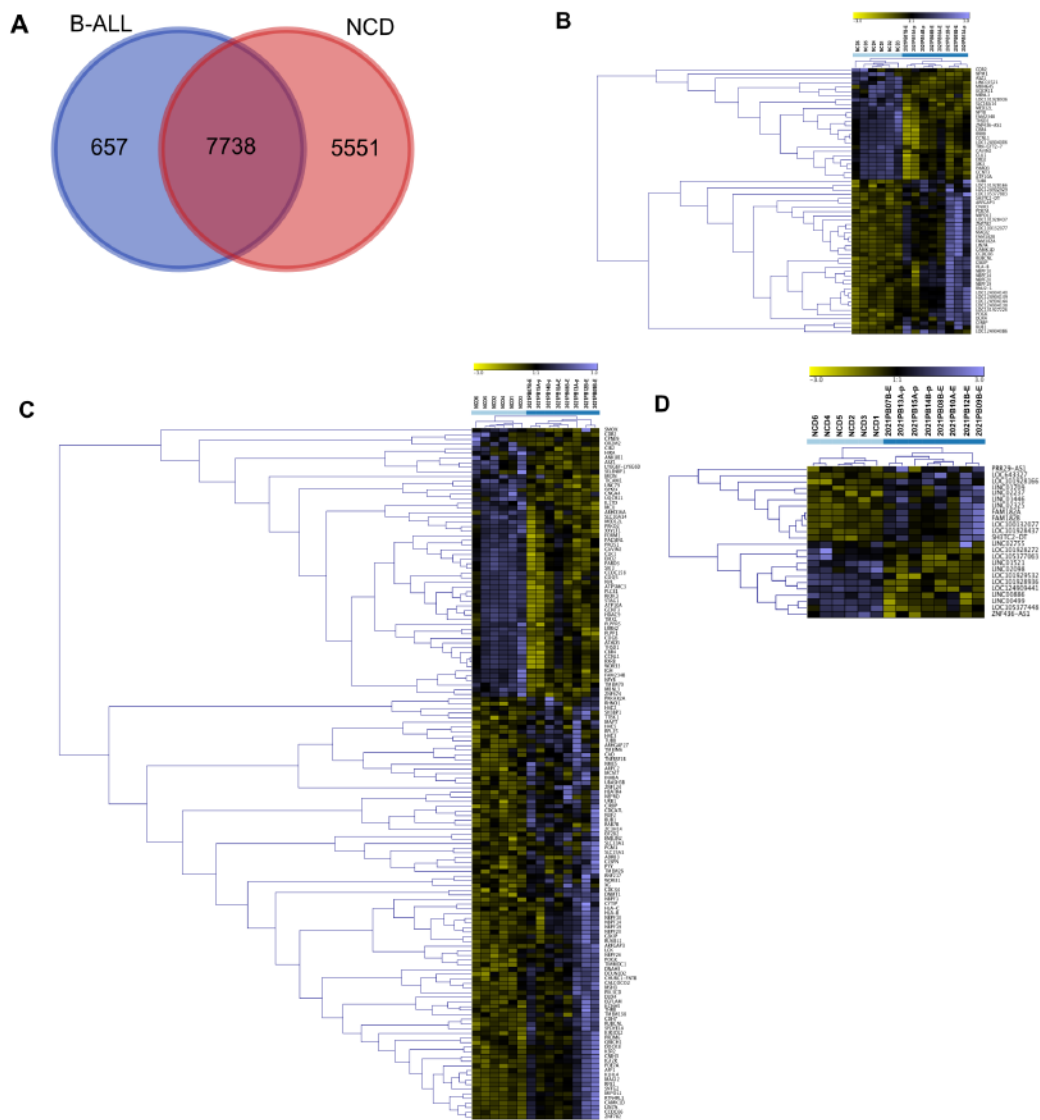


Fig. 4.4. Analysis of small RNA using DESeq2 distinguishes pediatric B-ALL from NCD. All DPRNAs were manually categorized and underwent unsupervised hierarchical cluster analysis based on Z-scores. (A) Venn diagram showing comparison of samples with reads in 8/8 pediatric B-ALL samples and 6/6 NCD, (B) unsupervised hierarchical cluster for RNA agnostic signature ($p \text{ adj} < 0.0001$), (C) unsupervised hierarchical cluster for mRNA, (D) unsupervised hierarchical cluster for lncRNA. Light blue squares are for NCD, and dark blue squares are for pediatric B-ALL.

similar. Conversely, the B-ALL clade showed four subclades. Genes with increased packaging in NCD EVs include MED12L, MIDN, ASZ1, and IGH mRNA while genes that are increased in B-ALL include NBPF-10, -14, -19 and 20, and HLA-B and -C (**Fig. 4.4C**). lncRNA fragments in the signature included long intergenic regions (lincs) such as LINC01521 and LINC02237, antisense regions such as ZNF436-AS1 and divergent transcripts such as SH3TC2-DT. The cluster of lncRNAs showed two main clades by sample, again with NCD clustering away from B-ALL (**Fig 4.4D**). DPspliceosomal RNA, DPsnRNA, DPmiR, DPtRNA, and unclassified RNAs do not discriminate NCD from pediatric B-ALL (**Supp. Fig 4.4**). Overall, the DPRNA subtypes have different potentials for discriminating B-ALL from NCD; however, the total small RNA signature is very clearly discriminatory between NCD and pediatric B-ALL (**Fig. 4.4B**).

We next identified genes exclusive to 8/8 B-ALL samples and found that 657 RNA transcripts have reads in all B-ALL samples, but not present in 6/6 NCD (Fig 5A and Supp file 1). Similarly, 5551 RNA transcripts are exclusive to all NCD samples (6/6) (**Fig 4.4A** and **Supp file 1**). Each group's top 100 RNA transcripts, sorted by average expression, were grouped and Z-scores calculated for hierarchical cluster analysis. As expected, the B-ALL exclusive and NCD exclusive transcripts show a clear discriminatory pattern between B-ALL and NCD (**Fig 4.5**). Thus, two sets of discriminatory EV RNA signatures were identified. One is a signature indicative of the normal control donor RNA profile, whereas the second signature consists of pediatric B-ALL exclusive RNAs (present in all 8 of our pediatric B-ALL samples) that represent the disease state.

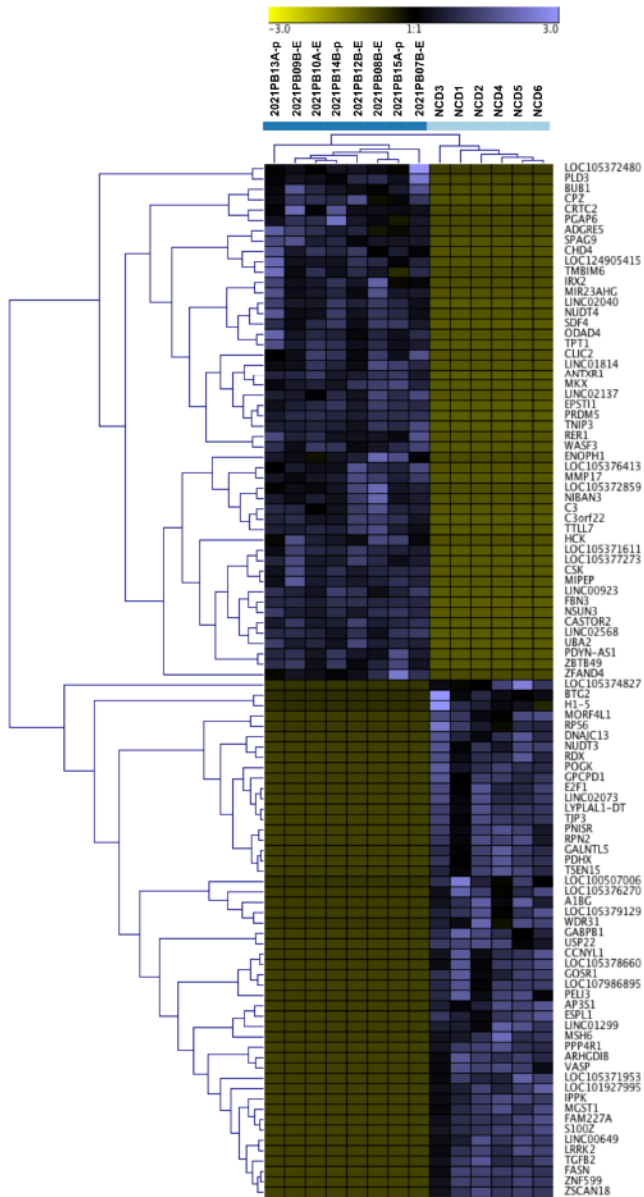


Fig. 4.5. Analysis of small RNA exclusive to B-ALL and NCD is highly discriminatory. Unsupervised hierarchical cluster of the z scores of relative log transformation of top 100 RNAs exclusive to pediatric B-ALL and NCD samples respectively. Light blue squares are for NCD, and dark blue squares is for pediatric B-ALL.

4.2.7 Both fragments and whole transcripts are packaged into EVs

Different small RNA species and regions were found to be packaged into the EVs. Only exonic regions are packaged into EVs for some genes, for example, DPmRNA MCM7 (**Fig 4.6A**). Unexpectedly, the introns of other genes were found to be packaged into EVs, such as DP lncRNA PRR39-AS1 (**Fig 4.6B**). MALAT1, a lncRNA of 8779 nucleotides, had multiple fragments of the single exon gene packaged into EVs (**Fig 4.6C**), and multiple miRNA genes on a single exon are packaged into EVs (**Fig 4.6D**). This data suggests that full transcripts of some RNA subtypes such as miRNA are packaged into EVs. However, the potential that whole transcripts are packaged needs to be explored further before more concrete conclusions can be reached.

4.2.8 Gene set enrichment analysis (GSEA) shows that EVs package gene sets that negatively regulate the cell cycle.

Using GSEA, we found that gene sets within the full signature (exclusive to B-ALL and DP RNAs) are preferentially packaged into EVs (activated). This gene set includes negative regulation of cell cycle processes, checkpoint signalling, response to stress, and cellular and biological processes (**Fig. 4.7A**). The genes associated with the identified gene sets are shown in relation to the associated gene sets (**Fig. 4.7B**). This illustrates that many of the genes are associated with processes related with the overarching theme of negative regulation of cell cycle (**Fig 4.7B**). Genes such as CLSPN and BUB1 are linked to negative regulation of cell phase transition, while MIR-193A is linked to negative regulation of mitotic cell cycle phase transition. Representative plots of specific gene sets exemplify the high enrichment scores for “negative regulation of cell cycle” and “negative regulation of mitotic cell cycle phase transition” (**Fig. 4.7C, D**). These data suggest that

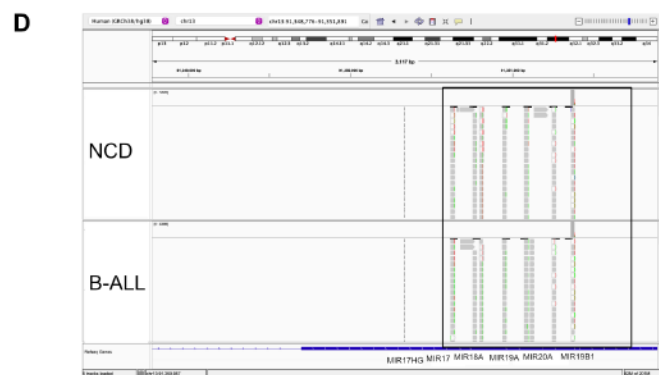
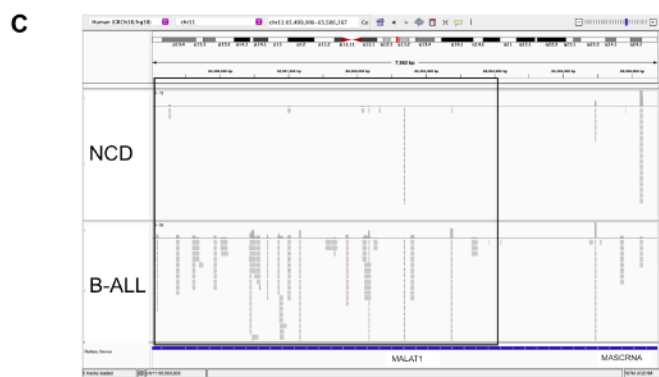
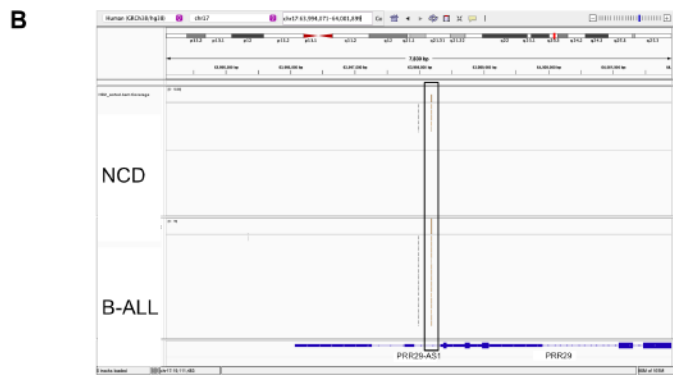
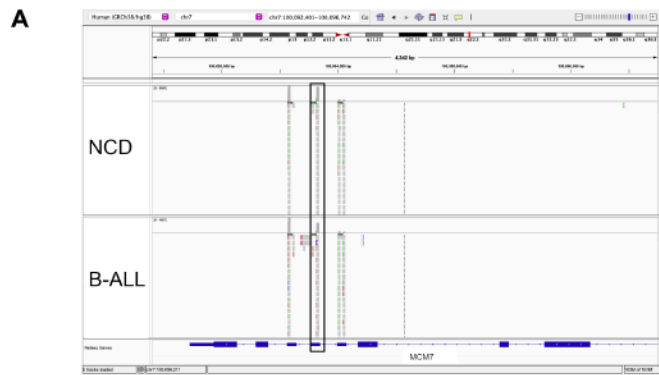


Fig. 4.6. Visualization of RNA-seq data from selected DPRNAs using an integrative genome viewer (IGV) showed different RNA packaging patterns. (A) A representative mRNA transcript, MCM7, with eight exons and eight intronic regions is shown. The black box highlights differential packaging of exon 2, with lower levels in B-ALL. (B) A representative lncRNA transcript, PRR29-AS1, contains five exons and four introns. The black box highlights a transcript from intronic region two as being differentially packaged between B-ALL and NCD, with higher levels in B-ALL. (C) A representative lncRNA transcript, MALAT1, is shown with a single exon. The black box highlights the multiple fragments of MALAT1 lncRNA which are differentially packaged between B-ALL and NCD EVs. (D) Several representative miRNAs (from left to right: miR-17, miR-18A, miR-19A, miR-20A, and miR-19B1) are encoded on a single continuous coding region on chromosome 13 and are packaged into NCD and B-ALL at relatively the same levels. Key: For each panel, the top toolbar shows the specific location address of the transcript and its relative position on a chromosome. The top portion of each track shows the relative abundance in log scale while the lower portion of each track show the specific fragment sequenced. Within each panel, the top section is a non-cancer donor (NCD, of each panel) and the bottom section is a pediatric B-ALL patient sample.

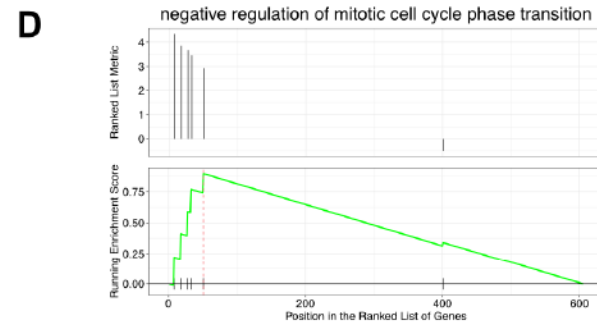
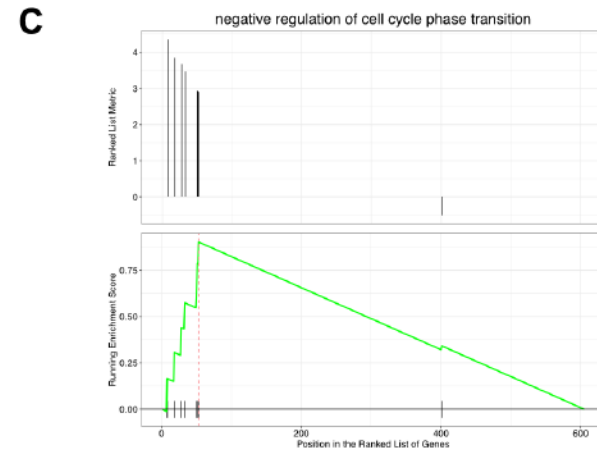
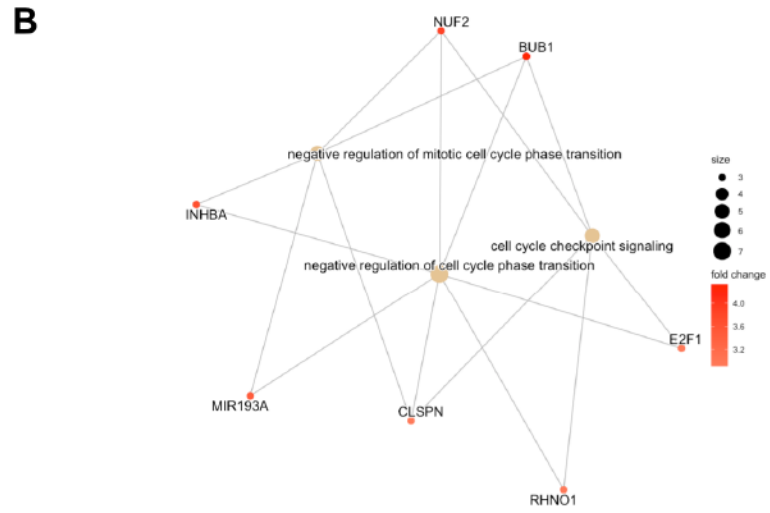
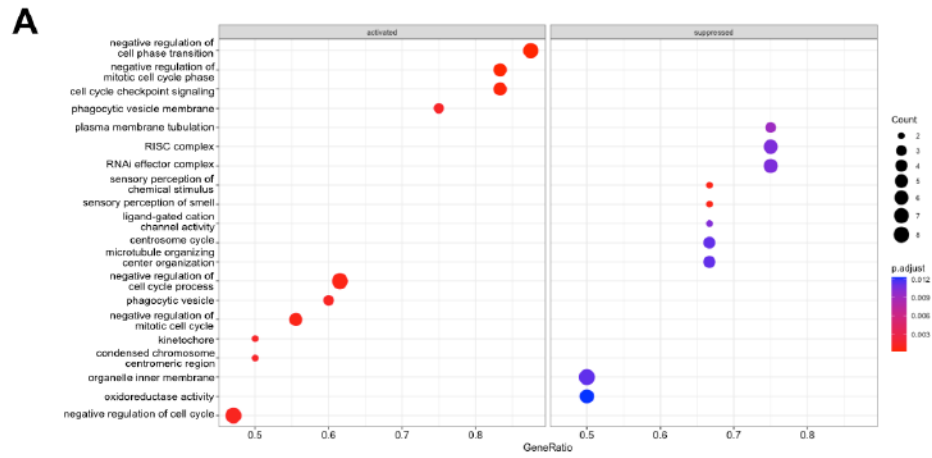


Fig. 4.7. Gene set enrichment analysis (GSEA) shows that transcripts that negatively regulate the cell cycle are enriched in B-ALL EVs (A) Enriched and suppressed pathways, with the gene ratios shown on the x-axis and the associated biological process on the y-axis. The left-hand side of the plot shows gene sets that are overrepresented in B-ALL EVs (activated pathways) and right-hand side of the plot shows genes sets that are excluded from B-ALL EVs (suppressed pathways). (B) enriched gene sets and known associated RNA species (orange circles – mRNAs or miRNA, yellow circles – gene sets) (C, D) Enrichment plots where the upper half is the ranked list metric, showing the ranking of the gene set in the list, and the lower half is the running enrichment score, including the position in the ranked list of genes. A total of 605 genes was annotated and analyzed.

the plasma EVs in B-ALL may preferentially package transcripts that impede proliferation, consistent with a role for EVs in the removal of unwanted transcripts.

4.2.9 GO and KEGG pathways of exclusive signature.

The GO and KEGG analysis of the top 200 NCD-exclusive EV RNAs showed key processes including semaphorin binding, intercellular bridge, interleukin 17-production and regulation of TORC1 signaling (**Fig 4.8A, Supp File 4.1**). Other pathways include protein maturation regulation, histone H3-K9 and neuromuscular junction development. Conversely, the top 200 pediatric B-ALL-exclusive EV RNAs showed pathways such as tumor-necrosis factor mediated signaling, choline metabolism in cancer, inositol phosphate metabolic process, platelet alpha granule lumen processes and protein deacetylation (**Fig 4.8B, Supp File 4.2**).

4.2.10 Competitive endogenous RNA (ceRNA) analysis shows miRNA: mRNA pairs.

To identify the ceRNA network associated with the total DPRNA signature, we identified miRNA: lncRNA and miRNA: lncRNA interactions using lncRRIsearch and targetscan. We identified miRNA: mRNA interactions using miRcode and miRwalk. We found lncRNA: mRNA, pairs instead of miRNA: lncRNA: mRNA triplicates (**Table 4.3**). This is not totally unexpected, as the ceRNA field is currently in its infancy, especially in pediatric B-ALL, and the interaction networks are based on what has been previously published.

The targets of the 4 DPmiRNAs from the RNA agnostic pipeline were predicted using two different platforms and overlapping target mRNA between both platforms used for visualization of miRNA: mRNA interaction networks (**Supp File 4.3 and 4.4**). Predicted mRNA targets were

compared to DPmRNAs (Fig 4.9), showing that there are 148 DPmRNAs, and 1999 mRNA targets, of which 14 mRNA are overlapping between the predicted mRNAs and packaged mRNAs. These 14 mRNAs are ATP10A, RNF217, MBNL3, MED12L, CDH7, MIPOL1, NFYB, AFF1, SLC15A1, MAGI2, KCNH5, XG, CDC34, KSR2. Overall, our data suggests that the different RNA transcripts differentially packaged into pediatric B-ALL EVs interact with each other in varying ways that are not yet fully understood.

4.3 Discussion

In this pilot study, we found that pediatric B-ALL peripheral blood plasma EVs have a similar size profile to NCD peripheral blood EVs, with both sample groups having similar mean and mode EV diameter. This average size range suggests that the EVs in both groups of samples could be larger exosomes and smaller microvesicles. With reference to EV numbers, the overall concentration of EVs was higher in pediatric B-ALL than NCD plasma. This finding aligns with what has previously been discovered in other cancers⁵³⁴⁻⁵³⁸, including in AML⁴¹⁵ and CLL⁴²⁴. We then used the clinically amenable Vn96 synthetic peptide to isolate EVs from plasma samples and subsequently, identified a signature of EV small RNAs that is exclusive to NCD and B-ALL, as well as RNA sequences that are differentially packaged. The small RNAs found in B-ALL EVs tend to be overrepresented with genes that negatively regulate the cell cycle, suggesting that EVs may be used to discard RNA sequences that are inhibitory to B-ALL growth and progression". In contrast, NCD EVs have a mix of different transcripts that appear to originate from multiple organs, including the central nervous system, muscle, platelets, and epithelial cells.

Consistent with our NTA findings, the Boyiadzis group quantified extracellular vesicles (by

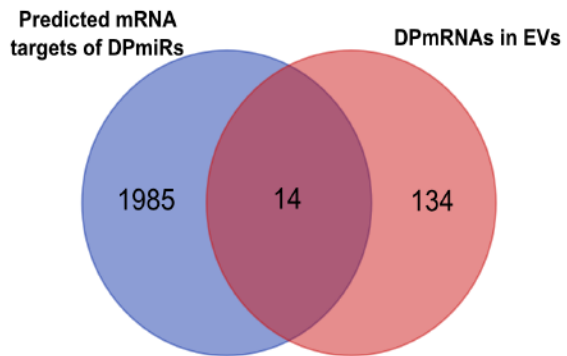


Fig. 4.9. miRNA and their targets are packaged into EVs, as shown in the Venn diagram showing mRNA targets of differentially packaged miRNA and mRNA cargo of EVs.

protein content) of sera from normal controls and AML patients, and found a higher EV concentration in AML EVs than normal controls ⁴¹⁵. Furthermore, targeted protein profiling of leukemia-associated antigens CD34 and CD117 in AML EVs versus control EVs showed differential representation of these markers, which were higher in AML than controls ⁴¹⁵. Using a different cohort of donors and samples, the Boyiadzis group used TGF- β 1 levels to track EV concentration, and found that AML EV burden is reduced post-chemotherapy induction, increased during consolidation, and normalized during complete remission, suggesting that AML EVs reflect leukemia burden ³⁷². Similarly, in B-cell CLL, the EV concentration was higher in patient blood plasma than in normal controls ³⁶⁵. Furthermore, CLL EV concentration is reduced post-therapy ⁴²⁴. Thus, plasma EV concentration may be a surrogate marker for monitoring B-ALL disease progression and response to therapy; however, this finding will require validation in a larger cohort of samples that includes post-treatment samples.

A key aspect of EV biology that determines their functionality is their cargo. Specifically, EV RNA cargo has been shown to be very important for overall EV functionality and interactions with recipient cells and environments ⁵³⁹. Furthermore, EV cargo has been shown to facilitate cancer progression and disease aetiology ⁵⁴⁰. Therefore, we decided to examine the small RNA cargo profile of pediatric B-ALL EVs, as a window into the disease. To identify the RNA profile and signature, we used the sRNADE platform (comprehensively called sRNAtoolbox) to annotate the RNA transcripts detected in the plasma EVs. Interestingly, both mature and precursor miRNA (pre-miRNA) transcripts were detected. While mature miRNAs have been widely reported to be packaged into EVs in different biofluids and diseases, including acute myeloid leukemia ³⁷⁶, pre-miRNAs have rarely been reported. However, tumour-derived microvesicles (TMVs) isolated from LOX melanoma cells carry pre-miRNAs such as pre-miR-21, pre-miR-27, pre-miR-100 and

pre-miR-151⁵⁴¹. Other transcript types that we also identified as packaged into EVs included non-coding RNAs (ncRNA) such as snRNA, snoRNAs, YRNAs and lncRNA. The distinction between intact versus fragmented was determined based on IGV visualization of the RNA transcripts, where we observed evidence for both. mirVana isolates total small RNA (<200 nt), but RNA transcript types such as lncRNAs and tRNAs are >200nt. Thus, the data suggests that both fragmented and intact mRNAs, tRNAs, and lncRNAs are packaged in EVs. The key observation that fragmented mRNA seems to be the largest component of our EV RNA cargo has opened the door to more questions. While it is not clearly known if this large amount of mRNA fragments is relevant or not, there may be a few explanations. This these mRNA may be a byproduct of increased gene expression in leukemia cells, as part of a mechanism for maintaining cell proliferation and survival. Conversely, these large number of fragments may be specifically packaged into EVs for leukemia relevant purposes such as waste disposal, shuttling out anti-cancer RNA or shuttling pro-cancer genes to distal locations to create a leukemia permissive environment for invasion of extramedullary sites. These speculations would have to be tested out in subsequent studies.

While this is the first time that these broad range of EV RNA cargo has been reported in pediatric B-ALL, other research groups have previously identified similar RNA species as part of EV cargo in healthy and diseased cells^{522,542-544}. Due to the often-pivotal role of different RNA transcripts in physiological conditions, differential expression of RNA transcripts in specific tissues is implicated in the aetiology of many diseases. Differential expression of mRNA, lncRNAs⁵⁴⁵, miRNAs⁵⁴⁶, snoRNA⁵⁴⁷, and circRNAs⁵⁴⁸ have been identified in pediatric B-ALL cells. Because RNAs have been shown extensively to be packaged from cells into EVs, it is plausible that some differentially expressed RNA transcripts in pediatric B-ALL could be packaged into EVs;

however, differential expression in cells does not automatically imply differential packaging into EVs.

The miRNAs in the miRNA-specific signature we identified using the sRNADE platform include miR-455-3p, miR-483-5p, miR-25-5p, miR-128-3p, miR-320b and -c, and miR-4485-3p, which we found are generally upregulated in pediatric B-ALL EVs in relation to NCD EVs; however, analysis using this platform did not show a clear miRNA-based discrimination between NCD and B-ALL. DPmiRs may act as tumor suppressors (TSmiRs) or oncogenes (oncomiRs). For instance, miR-455-3p functions as a tumor suppressor in pancreatic cancer ⁵⁴⁹, osteosarcoma ⁵⁵⁰, and non-small cell lung cancer ⁵⁵¹; miR-483-5p is downregulated in ovarian cancer cells and plasma ⁵⁵², while miR-25 is upregulated in AML cell lines and patient plasma ^{553,554}; miR-128-3p has been identified previously as a potential biomarker for ALL MRD ⁵⁵⁵ and as an oncomiR in T-ALL ⁵⁵⁶; miR-320b was found differentially expressed in CML ⁵⁵⁷, while miR-320c was identified in CLL as a p53 associated gene ⁵⁵⁸; miR-4485 is associated with myelodysplastic syndrome ⁵⁵⁹; however, whether these miRNAs act as tumor suppressors or oncomiRs in the pediatric B-ALL context is yet to be determined. Other DPRNAs identified using the sRNADE platform, such as miR-128-3p ^{560,561}, miR-24 ⁵⁶², and miR-181b-5p ⁵⁶³, have previously been identified to be upregulated in ALL. The rest of the DPmiRs have yet to be linked to pediatric B-ALL; however, in previous studies, miR-483-5p has been found upregulated in CLL blood plasma ⁵⁶⁴ and chronic myeloid leukemia (CML) cells ⁵⁶⁵. High miR-25 predicts a favourable outcome in acute myeloid leukemia AML ⁵⁵³, while high miR-1246 distinguishes control donor samples from AML patient samples ³⁷⁶. Some DPmiRs, such as miR-4665-5p, miR-758-3p and miR-744-3p, have been shown to be downregulated in leukemia ⁵⁶⁶. miR-3605-3p is upregulated in leukemia (ALL and CML) plasma samples as cell-free RNA (cfRNA) ⁵⁶⁷, while the others have not been previously found in

leukemia. In this present study, the types of DPmiRs identified by our analysis suggest that pediatric B-ALL cells may discard anti-tumour miRNA by packaging them into EVs, a mechanism to eliminate regulatory transcripts that could prevent disease progression.

The EV RNA sequences that are exclusive to all NCD samples include miR12136, a regulator of translation⁵⁶⁸, NOC2L, a regulator of histone acetyltransferase activity⁵⁶⁹ and PERM1, a regulator of mitochondrial content and oxidative function⁵⁷⁰. In addition, lncRNA LINC01134 directly binds chromatin which directly binds the promoter of AKT1SI, in turn leading to activation of the NFkB signalling pathway⁵⁷¹. Another prevalent lncRNA, LINC01714, is known to suppress proliferation, migration and invasion in cholangiocarcinoma (CCA) cells⁵⁷². Combined, NCD-exclusive EV RNA transcripts appear to be associated with generic and varied cell types. This suggests that in normalcy, blood cells and cells with access to the bloodstream secrete EVs and contribute to the total plasma EV repertoire. Conversely, the EV RNAs exclusive to B-ALL samples include C3, a complement protein, which is a serum effector of innate immunity. High C3 levels indicate inflammation, which may be associated with many hematological malignancies⁵⁷³. Other top genes include ODAD4, which is associated with ALL⁵⁷⁴, LINC02568, which is linked to squamous carcinoma⁵⁷⁵ and MRPS31, which is a driver of carcinoma⁵⁷⁶. Taken together, many of the B-ALL exclusive EV RNA transcripts are associated with disease, including hematological malignancies and immune activation.

Additionally, some of the DPmRNA in our signature are known to play active roles in leukemia and other cancers. mRNAs such as the LIN7A is differentially expressed in most samples and is known to be upregulated in AML⁵⁷⁷, and PDE7A, which is upregulated in most B-ALL samples in our study, is known to be upregulated in CLL and is associated with elevated cAMP levels⁵⁷⁸. Some known leukemia-associated transcripts such as DUX4⁵⁷⁹, IGH⁵⁸⁰, NFYB⁵⁸¹ and MED12L

⁵⁸² are notably differentially packaged into pediatric B-ALL, with DUX4 being upregulated and IGH, NFYB, and MED12L being downregulated in pediatric B-ALL EVs. Some of the mRNAs with higher levels in pediatric B-ALL EVs are known to be pro-tumorigenic. For instance, MCM7, a marker of a high proliferation rate in cancer, is upregulated in pediatric B-ALL EVs ⁵⁸³. CDC34, which is upregulated in pediatric B-ALL EVs, is known to facilitate lung cancer progression ⁵⁸⁴, while NUF2, which is upregulated in the EVs of 6 of 8 pediatric B-ALL samples in this study, is known to facilitate the progression of lung adenocarcinoma ⁵⁸⁵. Finally, HLA-B and -C are suggested to be involved in tumorigenesis of different non-leukemia cancer types ⁵⁸⁶. Potentially, these pro-tumorigenic mRNAs may be increasingly packaged from the lymphoblasts in the bone marrow into pediatric B-ALL EVs because of their overrepresentation in B-ALL lymphoblasts.

We found that the DPmiRs identified using the RNA agnostic DESeq2 platform varied from the DPmiRs identified using the miRNA-specific platform sRNADE. This may be because of differences in the gene count data, as each platform uses a different algorithm for count data processing, which may skew the total counts in the miRNA specific versus RNA agnostic platforms. The DPmiRs from the RNA agnostic platform are miR-193A, which is higher in EVs from most of the pediatric B-ALL samples, and miR-4645, miR-3143, and let7I, which are lower in EVs from pediatric B-ALL samples. miR-193A is upregulated in pediatric B-ALL patient blood plasma ⁵⁸⁷; miR-3143 is found to be downregulated in triple negative breast cancer ⁵⁸⁸ and is linked to B-ALL ^{589,590}; miR-let7I is associated with larynx and hypopharynx cancer ⁵⁹¹ and chronic myelogenous leukemia ⁵⁹². Taken together, the DPmiRs identified from the two different platforms (sRNADE and DESeq2) were not congruent or discriminatory enough for our purposes but matched previously known links to different cancers and in some cases, some types of leukemias, including B-ALL.

LincRNAs, which characteristically do not overlap with protein-coding transcripts, give rise to intergenic long non-coding transcripts⁵⁹³. So far, none of the lincRNAs we identified in this study are currently known to be involved with pediatric B-ALL or leukemia; however, some of the lincRNAs are linked to other cancers. For instance, linc00499 is known to be differentially methylated in glioblastoma⁵⁹⁴, while linc00886 is associated with ferroptosis in glioma⁵⁹⁵. Furthermore, linc01446 is higher in glioblastoma than in healthy donor cells⁵⁹⁶. Linc01521 is associated with sarcoma⁵⁹⁷. Linc02325 is part of a lincRNA signature for head and neck squamous cell carcinoma⁵⁹⁸. Linc02755 is associated with colorectal cancer as a predictor of metastasis⁵⁹⁹. Though a lot is not known about these lincRNAs in pediatric B-ALL, they may be potential targets for follow-up. Taken together, these DP-lincRNAs may need to be studied further to ascertain their relevance or lack thereof, in pediatric B-ALL.

We further explored the significance of the exclusively packaged and differentially packaged RNA in B-ALL biology using GSEA. The gene sets activated/enriched include negative regulation of cell cycle checkpoint signalling. It is currently known that dysregulation of the cell cycle is a hallmark of B-ALL⁶⁰⁰. The uncontrolled cell proliferation that results from cell cycle dysregulation could lead to leukemogenesis⁶⁰¹. Other gene sets activated/enriched in pediatric B-ALL EVs are phagocytic vesicle, kinetochore, and condensed chromosome centromeric region and organelle inner membrane. Overall, cell cycle and membrane-associated gene sets are activated (enriched). Conversely, gene sets such as membrane tubulation, RISC complex, oxidoreductase activity, and ligand-gated cation activity are suppressed (reduced). The GSEA is consistent with the hypothesis that packaging anti-tumour genes into EVs may drive an overactivation of pro-cancer pathways and processes to facilitate cancer tumorigenesis. Alternatively, the results of the GSEA may indicate increased expression of the DPRNAs in cells,

leading to increased packaging of these RNAs into EVs. Gene sets that are excluded from B-ALL EVs are associated with oxidoreductase activity, organelle and plasma membrane organization, regulation of transcription, and signal transduction. Developing B cells, which start as hematopoietic stem cells (HSCs), need cellular redox homeostasis to maintain a dynamic balance needed for their normal proliferation^{602,603}, which begins in the bone marrow (BM)⁶⁰⁴. The BM niche is hypoxic and generates reactive oxygen species (ROS) levels that are lower than in normal tissues, meaning that the BM niche is highly sensitive to changes in oxidative stress linked to oxidoreductase activity⁶⁰⁵. Changes in oxidoreductase levels can lead to alteration of ROS levels, which can affect the HSCs, including developing B cells⁶⁰⁶. Furthermore, moderate ROS acts as a second messenger that, in turn, regulates cell proliferation, while high ROS levels are linked to leukemic transformation^{607,608}. Thus, retention of these transcripts in the cell, as evidenced by reduced packaging in EVs, may promote the adaptation of the cells to the BM microenvironment. Taken together, the GSEA show themes that align with current knowledge about pediatric B-ALL onset, such as ligand-gated cation channel activity, cell cycle regulation and cell cycle phase transition^{609,610}.

In addition, some of the pathways associated with pediatric B-ALL exclusive genes such as ‘tumor necrosis factor mediated signaling pathway’ and ‘choline metabolism in cancer’ may suggest that there are cancer driver genes packaged into EVs. Conversely, passenger driver genes may also be present, as evidenced by pathways such as ‘negative regulation of epithelial cell apoptotic process’ and ‘sulfur compound transport’. This supports current knowledge about cancer, that driver genes – genes which when mutated confer cancer cells with significant growth advantage, while passenger genes have neutral impacts on cancer cell growth⁶¹¹. Specifically, B-ALL relevant driver genes such as ETV6, PAX5, ATF7IP, IKZF1, FLT3, SETD2 and RB1 (as shown by Ueno and

colleagues)⁶¹² were identified as packaged into pediatric B-ALL EVs. Other RNA transcripts may be pediatric B-ALL passenger genes, though this is not clearly known. Therefore, we can surmise that the differentially packaged and exclusivity signature genes in pediatric B-ALL EVs are RNA transcripts that reflect the genetic changes that drive B-ALL progression. Additionally, the non-driver genes may be passive passengers that are increased in leukemia cell associated increased biosynthesis and are thus packaged into EVs.

Predicted miRNA: mRNA and lncRNA: mRNA interactions suggest that the packaging of genes into EVs is non-random and coordinated in a way that is not yet fully understood. For instance, 14 of the mRNAs packaged into pediatric B-ALL EVs are predicted targets of the DPmiRs from the RNA agnostic signature. Let7I targets CDC34, while miR-193A targets SLC15A1 and KSR2. For example, Let7I is lower in B-ALL EVs, while its target mRNA CDC34 is higher in B-ALL EVs. This suggests that the higher levels of CDC34 packaged into B-ALL EVs may be due to lower levels of Let7I that are present in the producing cells. For miR-193A, which is upregulated in 3 of 8 B-ALL EV samples but downregulated in 5 of 8 B-ALL patients and all NCD, its target mRNAs SLC15A1 and KSR2 are overexpressed in 4 of 8 pediatric B-ALL patient EVs (for SLC15A1) and 3 of 8 pediatric B-ALL patient EVs (for KSR2) and downregulated in all NCD EVs. In this case, it could be a case of maintaining homeostatic levels of the mRNA by increasing miRNA levels. Furthermore, miR-3143 and its targets MBNL3, ATP10A, NFYB, AFF1, and XG; and miR-4645 and its targets MED12L and NFYB, CDH7, MAGI2, MIPOL1, KSR2, KCNH5 and AFF1 are all differentially packaged. While miR-3143 and miR-4645 have similar patterns (lower in B-ALL EVs), their targets trend differently. MIPOL1, MAGI2, AFF1, and KSR2 are higher in 4 of 8 pediatric B-ALL patients compared to NCD, while CDH7 and XG are higher in 5 of 8 B-ALL patients compared to NCD. Other targets ATP10A, MBNL3, MED12L are lower in 7 of 8 B-ALL

patients compared to all NCD, while NFYB is lower in all B-ALL patients compared to NCD. Therefore, the DPmiRs and their target DPmRNAs trend without clear consistent patterns. Taken together, the DPmiR and target DPmRNA packaging into pediatric B-ALL EVs suggests the packaging of genes important in disease onset and progression.

The implication of EV RNA cargo with such diversity is far reaching for disease aetiology. Our KEGG, GSEA and ceRNA analysis suggests that these EV RNA cargos are not random and may be associated with aspects of leukemia onset and progression. Leukemia onset and progression is highly linked to the tumor microenvironment which contains stroma and non-stroma cells, with EVs potentially acting as conduits for transfer of RNA and protein cargo to recipient cells where pro-cancer changes can be made. Johnson and colleagues showed that ALL derived EVs when taken up by normal stromal cells altered normoxic metabolism leading to decreased mitochondrial respiration and increased extracellular acidification linked to lactate production – a reverse Warburg effect^{382,613}. The specific cargo is ALL derived EV that mediates this change upon uptake by stromal cells is currently unknown. However, in CLL, it was shown that CLL derived EVs are taken up by stromal cells, resulting in transfer of miR-202-3p which modulates Sufu expression, leading to reprogramming the stromal cells phenotypic⁶¹⁴. Furthermore, CLL derived EVs transfer phospho-receptor tyrosine kinase Axl to BMSCs, which activate AKT/mTOR/P70S6K/hypoxia-inducible factor-1 alpha axis resulting in the production of vascular growth factor, a survival factor for CLL B cells³⁸⁴. Huan and colleagues showed that uptake of AML derived EVs by stromal cells resulted in transfer of CXCR4 mRNA, which is suggested to reprogram the microenvironment to be more pervasive to invasion of the BM by AML³⁷⁴. In addition, FLT3 and CXCR4 was transferred from AML derived EVs to stromal cells³⁷⁷. Chronic myelogenous leukemia (CML) derived EVs transfer amphiregulin to BM stromal cells upon EV uptake leading to activation of

epidermal growth factors signaling and increase in expression of SNAIL and its targets MMP9 and IL8³⁸⁰. Therefore, it can be surmised that pediatric B-ALL EVs may be acting along the lines of transferring cargo to recipients in a manner that facilitates disease progression. However, the specifics of how this mechanism works will need to be explored in follow-up studies.

B-ALL is generally a heterogenous disease, with immunophenotypically heterogenous cell populations and distinct subpopulations having broad expression of B-ALL hallmark markers coexisting at once. For instance, the presence of multiple disease subclones are reported in B-ALL, even as early as at diagnosis⁶¹⁵. Therefore, interpretations of our results cannot only be viewed in the light of pediatric B-ALL subtypes, as each patient sample contains contributions from multiple B-ALL subclones.

Despite our best efforts, there are a few limitations of this study. The first limitation is the type of control samples. While it would be preferable to get healthy age and sex matched donors as healthy controls, it is rare to find healthy children becoming sample donors for any biofluid. In the face of that, we resorted to age matched donors who did not have pediatric leukemia, hence non-cancer donors. These donors had a range of non-cancer related disorders such as tonsillitis. Thus, it is possible that they are a slight deviation from completely healthy but were sufficient since they were not leukemia patients. The second limitation the small sample size (6 non cancer donors and 8 pediatric B-ALL patients (5 ETV6::RUNX1 subtype and 3 BCR::ABL -like subtype), which may affect statistical significance. The solution to determine if this study was sufficiently powered, was to conduct a power analysis. At the time of designing the experiments, we attempted to run power analysis for RNA-Seq on EVs samples using the method postulated by Wu and colleagues⁶¹⁶. However, we were limited by the design of the tool, which is not yet designed to

input custom experimental details. However, this is a proof-of-concept, which will be further validated using a larger validation cohort.

4.4 Conclusion

EVs play a role in shuttling signalling mediators between cells. In this study, analysis of the small RNA cargo of pediatric B-ALL EVs suggests that packaging of this cargo is not random and may be involved in aspects of disease onset and progression. We also showed the presence of a signature exclusive to all the NCD samples (normalcy), and a disease-associated signature that is exclusive to all B-ALL samples. Pediatric B-ALL EV concentration is also a potential biomarker for pediatric B-ALL, with increased EV levels that may precede the increase of cells in circulation. In this way, EV concentration could be a less invasive surrogate biomarker for monitoring disease burden, therapy response, and even early detection of MRD. The potential for using the gene signature for either detection of disease burden or relapse monitoring or using EV concentration as a biomarker will need to be tested using a larger validation cohort. Moving forward, a study to validate/explore the findings from this study, using a larger patient cohort, with samples collected longitudinally and cross-sectionally will be the next direction. This will allow for exploring the results shown here to facilitate stronger conclusions based on the data.

Chapter 5: Small RNA transcripts are selectively packaged into pediatric acute lymphoblastic leukemia-derived extracellular vesicles.

5.1 Introduction

Globally, leukemia is the most prevalent pediatric cancer, accounting for one-third of pediatric cancer diagnoses in Canada⁶¹⁷. Acute lymphoblastic leukemia (ALL) is a leukemia of aberrant immature lymphocytes called lymphoblasts which differ significantly from normal maturing B-cell precursors and proliferate within the bone marrow, spreading to blood and extramedullary sites²¹². ALL is the second most common type of acute leukemia in adults and affects 80% of children, with 80% of ALL cases arising from B cell lineage (B-ALL)^{212,505}. Hence B-ALL represents a significant global cancer burden.

Extracellular vesicles (EVs) are lipid membrane-enclosed nanoparticles containing nucleic acid and protein cargo and are ubiquitously released by all cells tested to date²⁴⁶. EVs are classified based on their size and method of biogenesis into three subtypes (exosomes, microvesicles and apoptotic bodies). It has been shown that cancer cells can secrete up to 10^4 EVs per cell^{424,535} that can regulate gene expression in recipient cells in a pro-oncogenic manner⁶¹⁸. Notably, B-ALL blasts release a population of EVs that can reprogram the bone marrow microenvironment³⁸², facilitate cell-cell communication^{382,613}, expedite the transfer of pro-cancer cargo⁶¹⁹, and possibly play a role in drug resistance⁶²⁰. Furthermore, EVs are stable under various storage conditions⁶²¹, can cross the blood-brain barrier (BBB)¹⁷ and are abundantly detectable in all tested body fluids, including blood, urine, and cerebrospinal fluid²⁴⁴. Published studies on chronic and acute myelogenous leukemia, have demonstrated an increase in leukemia-specific EVs, which carry disease-specific protein and RNA markers⁶²⁴⁻⁶²⁷. Furthermore, we have observed an overall

increase in the particle number of EVs in the plasma of pediatric B-ALL versus non-cancer donor (NCD) samples (unpublished observations).

The nucleic acid cargo of EVs is distinct from that of the originating cell, while being representative of a subset of the originating cell's nucleic acid repertoire⁶²⁸. A variety of nucleic acids have been shown to be packaged into EVs including messenger RNA (mRNA), long non-coding RNA (lncRNA) and microRNA (miRNA)⁶²⁸. EVs are particularly enriched in miRNA, where a subset of cellular miRNA is packaged into EVs^{522,629,630}. The cellular mechanisms underpinning the packaging of miRNA and other RNA biotypes into EVs is still not fully understood.

There are two general processes involved in the packaging of miRNAs into EVs: loading and sorting. Loading is the recruitment of miRNAs to the site of EV biogenesis, resulting in the miRNAs being constitutively included in the EVs⁶³¹. Conversely, sorting is an active process in which miRNAs are selectively incorporated into EVs⁶³¹. It is now known that both processes occur in tandem^{380,631,632}. Loaded miRNAs may be those at the same level in originating cells and EVs, while sorted miRNAs will be at different levels in cells and EVs (low in cells and high in EVs or *vice versa*). For the lncRNAs, a selective packaging mechanism that may depend on active or passive processes has been elucidated in colorectal⁶³³ and prostate cancer⁶³⁴. Recipient cells can take up the lncRNAs-containing EVs, where the lncRNAs change the recipient cells' phenotype to facilitate cell proliferation, metastasis, angiogenesis and immunosuppression⁶²⁸. Different studies have reported lncRNA fragments^{635,636}, full length^{628,637} or both^{633,638,639} were detected in EVs. tRNAs, called 5'tRNA halves because they include the anticodon loop⁶⁴⁰, are also found in EVs. The packaging pattern of tRNA fragments points towards concentration-dependent selective packaging⁶⁴¹. snoRNA was reported to be found in the EVs isolated from plasma of Alzheimer's

disease patients and cognitively normal controls⁶⁴² and the peripheral blood of astronauts⁶⁴³. Not much is known about how snoRNAs are packaged into EVs or if they are differentially packaged into cancer cell derived EVs. snRNAs are packaged into the EVs of malignant melanoma cell lines A375, MML-1 and SK-MEL⁶⁴⁴.

Selective or non-selective sorting can occur in a cell context-specific manner^{297,645}. These context specific packaging patterns is dependent upon four main mechanisms, as outlined below. The first mechanism is RNA binding protein (RBP)-dependent. Studies have identified cell-specific seed sequence motifs within the sequence of EV-packaged miRNA, which are recognized by specific RBPs for selective sorting into EVs^{131,295,298,646-651}. The first type of RBP dependent pathway is the miRNA motif and SUMOylated heterogenous nuclear nucleoproteins (hnRNPs)-dependent pathway⁶⁵². In T lymphocytes, 'GGAG' seed sequence-containing miRNAs are recognized by RBP heterogenous nuclear ribonucleoprotein A2B1 (hnRNPA2B1) in a SUMOylation-dependent manner and selectively sorted into EVs⁶⁴⁷. In nontumorigenic mouse hepatocytes, EV-enriched miRNAs contained 'GGCU' extra-seed motifs that are recognized by RBP synaptotagmin-binding cytoplasmic RNA-interacting protein (SYNCRIP)⁶⁴⁸. Human embryonic kidney cells (HEK293T) derived EVs were enriched in miR223 and miR144, which were recognized by Y box protein I (YBX1)^{298,651}. The second mechanism is the miRNA 3' -end sequence-dependent pathway⁶⁵². Specifically, in lymphoblastoid B cell, Burkitt's Lymphoma, diffuse large B-cell lymphoma and human urine-derived EVs, EV-enriched miRNAs were uridylylated at the 3' end, while cell retained miRNAs were adenylated at the 3' end⁶⁵³. In addition, breast cancer cell (MDA-MB-231) derived EVs were enriched in miR122, which contained a 3' end 'UUU' sequence and/or UCGA motifs recognized by the Lupus La protein²⁹⁸. A third mechanism for miRNA sorting is miRNA induced silencing complex (miRISC)-related pathway⁶⁵². Argonaut proteins such as Ago2 are involved in

the RISC complex and in miRNA sorting into EVs for some cell types but not in others^{649,651}. Finally, sorting of miRNAs can depend on the neutral sphingomyelinase 2 (nSMase)-dependent pathway, via the synthesis of ceramide⁶⁵². Specifically, it was discovered by Kosaka and colleagues and Mittelbrunn and colleagues that inhibition of *de novo* ceramide synthesis via nSMase inhibition resulted in the reduction of miRNAs carried in exosomes^{654,655}.

Additionally, in the studies where RNA sorting mechanisms were identified, there seemed to be an overlap of mechanisms. For instance, Ago2 is both an RBP and a miRISC associated protein, which means some cells may utilize multiple mechanisms synergistically. Finally, it should be noted that RBPs (HNRNPs, Ago2, YBX1, MEX3C, MVP and La protein) and membrane proteins (Caveolin-1 and nSMase2) are involved in miRNA sorting into EVs from different cell types^{656,657}. Currently, it is not known whether identified packaging mechanisms are cell-specific, or if the motifs are recognized exclusively by specific RBPs or by a host of RBPs.

For the other RNA species which have been shown to be packaged into EVs, the packaging mechanisms are currently understudied. RNA species such as lncRNAs (>200nt), snRNA (~150nt), tRNA derived small RNAs (30 – 45 nt tRNA halves, 14 -30nt tRNA fragments, 30-35 nt 5'tiRNAs and 40 – 45 nt 3' tiRNAs), snoRNA (60 – 300nt), yRNAs (~120nt), and mRNA (>1000nt), have been shown to be packaged into EVs^{298,632,651,658,659}. However, to the best of our knowledge, the packaging of RNA into pediatric B-ALL EVs has not been explored. Here, we show for the first time that in addition to miRNAs, other RNA transcripts including lncRNAs, mRNAs, snRNAs and tRNAs are loaded and sorted into pediatric B-ALL EVs. Specifically, we explore RNA cargo from both primary pediatric B-ALL patient lymphoblasts (n = 6) and immortalized cell lines (n = 2) into EVs and show that these cells load (non-selectively) and sort (selective) multiple RNA biotypes into their EVs.

5.2 Results

5.2.1 Pediatric B-ALL lymphoblast and cell line EVs have similar size and concentration.

We first characterized the EVs secreted by the cell lines and primary lymphoblasts. The primary lymphoblasts were obtained from ETV6::RUNX1 and BCR::ABL-like subtype patients respective. The immortalized pediatric B-ALL cell lines are RCH-ACV and UoC-B1. RCH-ACV cells from an 8-year-old female with relapsed BM pre-B-ALL with hypodiploid karyotype and TCF::PBX1 mutations. UoC-B1 cell line is from a 15.5-year-old female with TCF3::HLF translocation. Using TEM, we found that B-ALL cell lines and primary lymphoblasts secrete round EVs with a double membrane (**Fig 5.1A**). Concurrently, NTA analysis showed the presence of a substantial number of nanoparticles that is similar between cell lines and primary lymphoblast cultures (**Fig 5.1B**). To validate that Vn96-based EV isolation isolated EVs, we detected the presence of EV markers (HSC70, HSP90, CD63, CD81), B cell-specific marker (BLNK) and purity marker (Calnexin) in immortalized pediatric B-ALL cell lines and their respective Vn96-isolated EVs (**Fig 5.1C**). We found that Vn96 isolated EVs, which are HSC70⁺, HSP90⁺, CD63⁺, CD81⁺, BLNK⁺ and Calnexin^{-/low}.

5.2.2 Different types of RNA transcripts are packaged into EVs.

Using sRNAtoolbox, different types of RNA species were identified in cell lines and primary lymphoblasts compared to isolated EVs. Between 10 – 40% of the transcripts were unassigned (**Supp. fig. 5.1**). The most abundant RNA species of the currently assigned transcripts are non-coding RNAs with little information (other RNA), so they cannot be properly classified (**Fig 5.2A**).

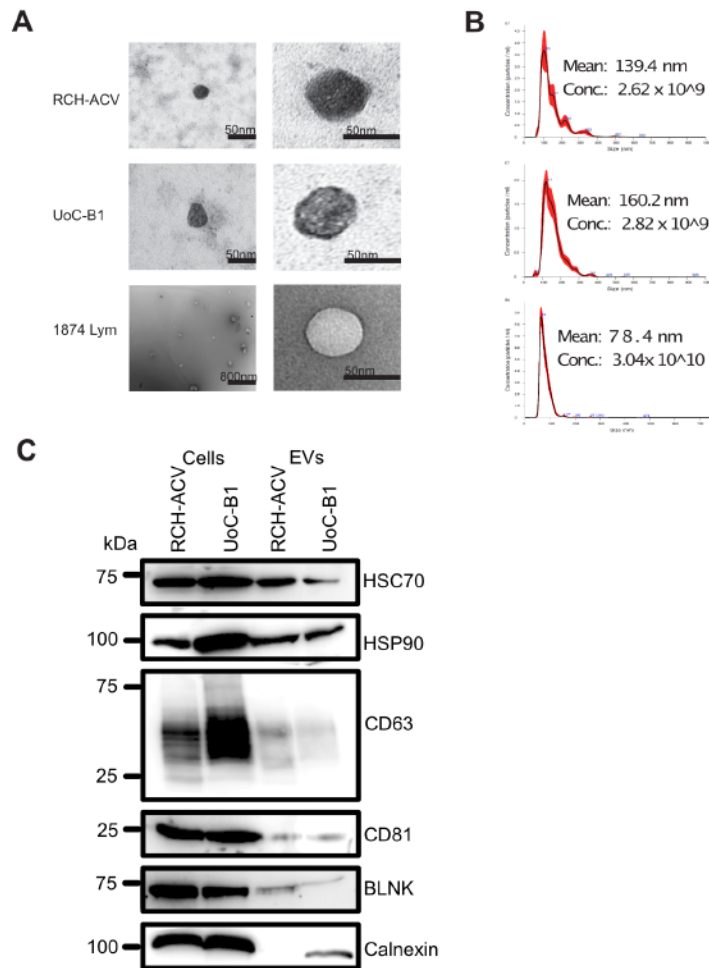


Fig. 5.1. Vn96-isolated EVs and conditioned media contains EVs with typical characteristics

(A) Transmission electron microscopy of negatively stained EVs from Vn96 pulldowns from pediatric B-ALL cell lines (RCH-ACV and UoC-B1) and cell conditioned media EVs from patient peripheral blood lymphoblasts. (B) Nanoparticle tracking analysis on EVs in cell conditioned media of RCH-ACV, UoC-B1 and patient peripheral blood lymphoblasts. (C) Western blot analysis of proteins from RCH-ACV and UoC-B1 cell lysates (columns 1 and 2), and Vn96 pull down EVs (columns 3 and 4). The location of molecular weight markers is shown on the left.

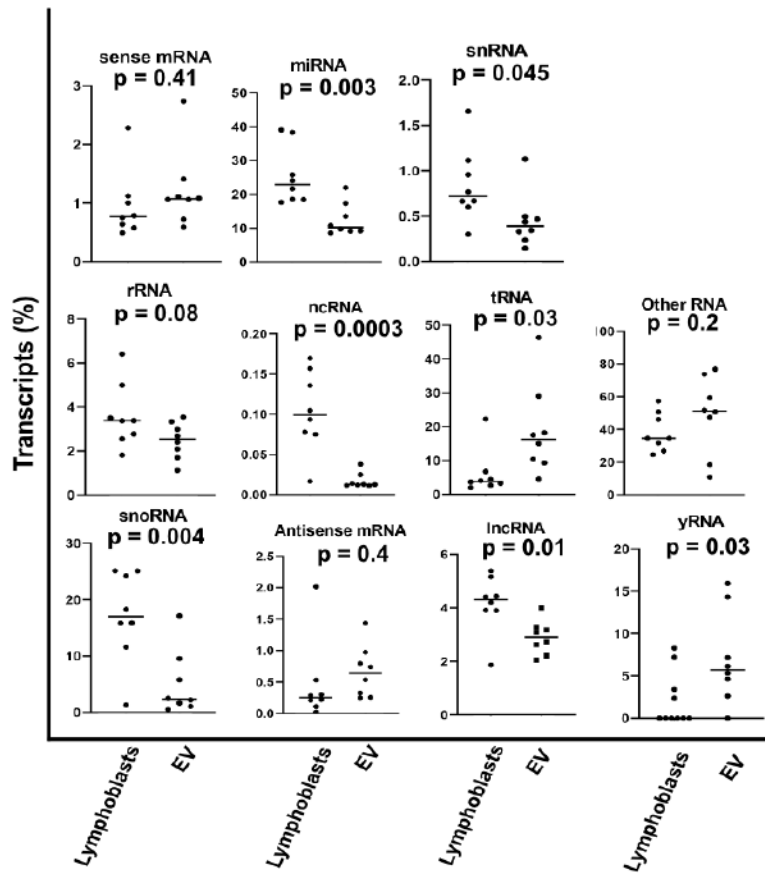
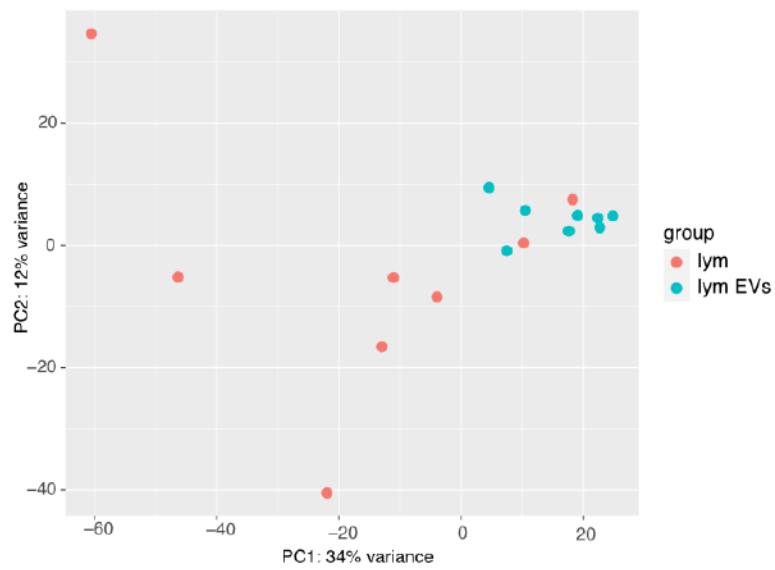
A**B**

Fig. 5.2. Overall transcript sub-type abundance differs in EVs compared to the originating lymphoblasts or cell lines. (A) SRNAde analysis of small RNASeq data comparing RNA species in cells versus EVs showed differential presence of RNA species in cells and packaged into EVs (Unpaired t test, $P < 0.05$). (B) principal component analysis of cells (primary lymphoblasts and immortalized cell lines) (lym, orange) and lymphoblast EVs (lym EVs, teal).

Overall, the next most abundant species in descending order are miRNAs, tRNAs, snoRNAs, yRNAs, rRNAs, lncRNAs, sense and antisense mRNAs, snRNAs, and lastly ncRNAs (**Fig 5.2A**). Of the detected RNA transcripts, miRNAs, snRNAs, rRNA, ncRNAs, snoRNAs and lncRNAs transcripts are more abundant in cells than in EVs. Conversely, sense and antisense mRNAs, tRNAs, other RNAs, and yRNAs are higher in EVs than in cells. Principal component analysis (PCA) showed that all EVs clustered together while cells did not all cluster together. This shows that pediatric B-ALL cells tend to package similar RNA cargo into their EVs making the EVs relatively similar (**Fig 5.2B**).

5.2.3 RNA transcript packaging mechanisms

We identified 7385 transcripts that were present in all cells (i.e. 8/8), and 8834 present in all the EV samples (**Fig. 5.3A**). Of these RNAs, 1220 RNAs were found to be cell exclusive while 2669 RNAs were EV exclusive (**Fig. 5.3A** and **Supp. file 5.1**).

For selective packaging, the top 100 cell exclusive RNAs and top 100 EV exclusive, based on average expression, were pooled and z scores calculated. These pooled RNAs were clustered using unsupervised hierarchical clustering (**Fig. 5.3B**). All cells, except for RCH-ACV, clustered closely together while EVs from both primary lymphoblasts and cell lines clustered together.

Using integrative genome viewer (IGV), we visualized different types of RNA packaging from lymphoblasts into EVs. H2BC7 (**Fig. 5.4A**) and miR-6724-4 (**Fig. 5.4B**) are examples of a cell exclusive transcripts while TRABD has EV exclusive and EV enriched fragments (**Fig. 5.4C**). LINC00189 contains an EV enriched fragment from an intron (**Fig. 5.4D**). In addition, RNAs such as RNA45SN1, RNA5-8SN1 and RNA28SN1 (**Fig 5.4E**) seem to have almost the full transcripts

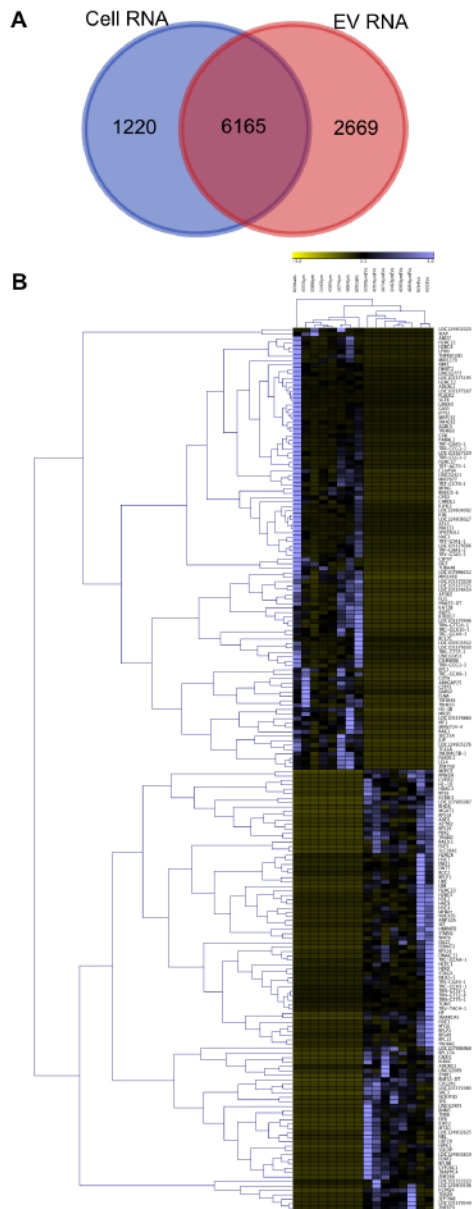


Fig. 5.3. Exclusively sorted RNAs are discriminatory between cells and EVs as shown by(A) Venn diagram of RNAs showing transcripts detected in all cells (primary lymphoblasts and immortalized cell lines) and all EVs derived from these cells. (B) Unsupervised hierarchical cluster of the top 100 cell exclusive and 100 EV exclusive RNAs by average expression level.

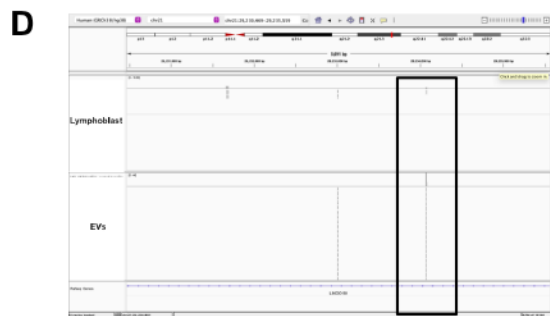
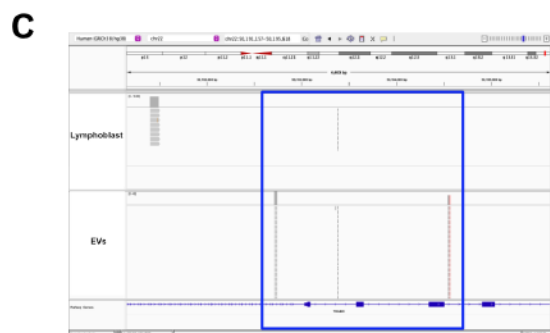
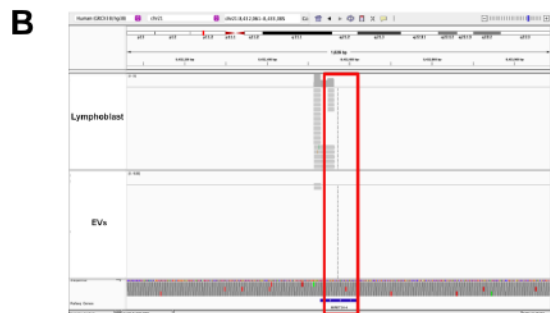
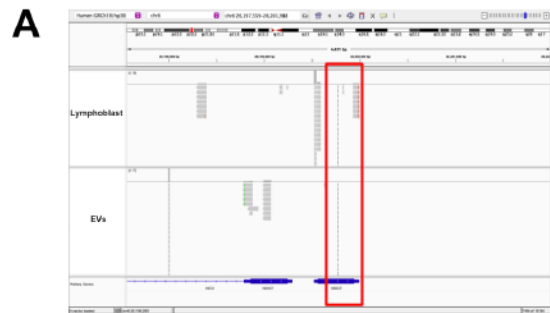


Fig. 5.4. Fragments of introns and exons as well as full length RNAs are packaged into EVs or retained in cells. (A) An example of a packaged fragment is from the single exon of H2BC7 mRNA, which is exclusive to cells. (B) An example of a miRNA (miR-6724-4) packaged in a cell exclusive manner. (C) An example of a fragment from one exon of the multi-exonic mRNA (TRABD), which is exclusive to EVs. (D) Within the intron of LINC01899, a fragment is shown to be present in cells and EVs, but higher in cells than EVs (E) Three linked rRNAs (RNA45SN1, RNA5-8SN1 and RNA28SN1) seem to have full-length RNAs packaged into EVs.

packaged into EVs, in contrast to other RNA types shown previously, which appear to be fragments.

5.2.4 Non-selectively packaged RNAs are enriched in cellular proliferation and metabolism.

The top 200 non-selectively packaged RNA transcripts were analyzed using GO and KEGG pathways in Cytoscape (**Supp. fig. 5.2**). The network analyses show that the non-selectively packaged RNAs are associated with pathways such as ‘electron transportation’, ‘anchored component of plasma membrane’, ‘ribosome biosynthesis in eukaryotes’, ‘positive regulation of organelle assembly’, and ‘mitotic sister chromatid cohesion’. These are pathways associated with normal cellular biology needed for maintaining physiology and division. Furthermore, pathways such as ‘protein kinase A binding’, and ‘oxidative phosphorylation’ are also associated with maintaining cell physiology in eukaryotic cells including B cells. However, some of the pathways are not associated with B cell physiology directly. These include pathways for ‘cardiac muscle contraction’, ‘folate biogenesis’, ‘columnar/cuboidal epithelial cell development’, ‘Parkinson disease’, ‘diabetic cardiomyopathy’, ‘Huntington disease’ and ‘prion disease’. It is unclear what the association between B cells or B-ALL is with these pathways.

5.2.5 Cell exclusive transcripts showed cell maintenance associated gene sets and pathways while EV exclusive transcripts include pediatric B-ALL and immune associated genes.

GO and KEGG network analysis by Cytoscape showed that the cell exclusive RNAs are associated with general cell maintenance genes such as for ‘negative regulation of actin filament polymerization’, ‘positive regulation of nucleocytoplasmic transport’, ‘ubiquitin dependent ERAD pathway’, and ‘phospholipid scramblase activity’ (**Fig. 5.5A**). In addition, ‘acetylglucosaminyltransferase activity’, ‘cellular response to zinc iron’ and ‘substrate adhesio

Fig. 5.5. Genes involved in cellular metabolism retained in cells while genes involved in immune function are packaged into EVs.

Network analysis of gene ontology (GO) and KEGG pathways for top 200 (A) cell exclusive and (B) EV exclusive RNAs. Coloured circles and labels are the main gene ontologies, while non-coloured circles and labels are subsidiaries, associated with the main genes.

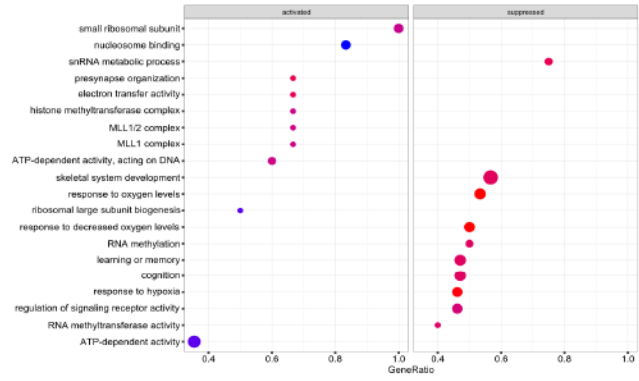
depending on cell spreading' pathways are also identified. These pathways suggests that cell exclusive RNA transcripts are core transcripts necessary for maintaining cell status and physiology. Furthermore, there was an abundance of transcription and translation associated pathways such as for 'nucleosome', 'H4 histone acetyltransferase complex' 'post-translational protein modification', 'cytoplasmic sequestration of proteins', 'regulation of protein maturation', and 'protein location to chromosome centromeric region'. These themes suggest that the cell exclusive RNAs may also be linked to maintaining cell function via regulation of gene expression at multiple levels. Furthermore, some signaling pathways were identified such as for 'negative regulation of transforming growth factor beta receptor signaling pathway' and 'regulation of insulin receptor signaling pathway'. These are signaling pathways known to be crucial for B cell development, and homeostasis. Other pathways such as 'embryonic heart tube morphogenesis', 'oocyte maturation', 'synapse maturation', 'blastocyst maturation', and 'megakaryocyte differentiation' are related to cell differentiation and development.

Network analysis of EV exclusive genes revealed key GO terms such as 'mRNA 3'UTR AU-rich region binding', 'chromosome centromeric core domain', 'nucleosome' 'DNA replication-dependent chromatin assembly', and 'positive regulation of transcription by RNA polymerase III' (**Fig 5.5B**). Furthermore, transcription and translation associated pathways such as 'positive regulation of translation', and 'protein localization to chromosome were also identified. This suggests that EV exclusive transcripts can affect gene regulation. Furthermore, other terms exclusive to EVs are immune response-related terms, such as 'leukocyte trans endothelial migration', 'negative chemotaxis', 'erythrocyte homeostasis', 'positive regulation of response to cytokine stimulus', 'negative regulation of GPCR', 'regulation of macrophage activation' and 'regulation of $\alpha\beta$ T cell differentiation'. This suggests that RNAs associated with B cell function

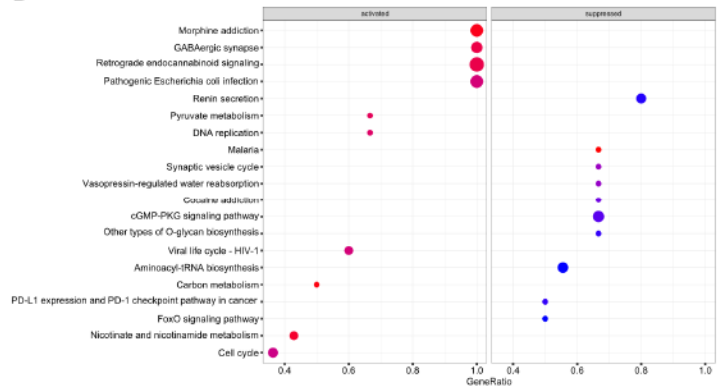
in the adaptive immune response are packaged into the EVs. Furthermore, pathways associated with infection including ‘bacterial invasion of epithelial cells’, ‘negative regulation by host of viral process’, ‘viral carcinogenesis’, ‘shigellosis’ and ‘Yersinia infection’ were also identified. Other immune-associated pathways include ‘negative regulation of IL-6 production’, ‘negative regulation of interferon-gamma production’, and ‘regulation of G protein-coupled receptor signaling pathway’ and ‘TGF-beta signaling’. These immune and infection-associated pathways further strengthen the point about immune system associated transcripts been packaged into pediatric B-ALL in an exclusive manner. Furthermore, the ‘MLL1/2 complex’ pathway is a B cell disease specific pathway. Thus, genes that regulate the immune response in general and pediatric B-ALL specifically are packaged into the EVs.

The gene set enrichment analysis (GSEA) for cell exclusive RNA transcripts revealed activation of GO terms including ‘small ribosomal subunit’, ‘nucleosome binding’, ‘presynapse organization’, ‘electron transfer activity’, ‘histone methyltransferase complex’, ‘MLL1/2 complex’, and ‘ATP-dependent activity’ (**Fig 5.6A**). Similar to the network analysis, this suggests that cell exclusive RNA activates pathways associated with cell cycle maintenance and gene expression. Also, some leukemia associated GO such as MLL1/2 complex are activated in the cell exclusive RNA. The suppressed GO include ‘snRNA metabolic process’, ‘skeletal system development’, ‘RNA methylation’, ‘response to hypoxia’, and ‘regulation of signaling receptor activity’ (**Fig 5.6A**). This data suggests that cell exclusive RNA transcripts are associated with response to metabolism, oxygen levels, and signaling. Activated pathways in the GSEA analysis include ‘pyruvate metabolism’, ‘DNA replication’, ‘HIV-1 viral life cycle’, ‘carbon metabolism’, and ‘cell cycle’ (**Fig. 5.6B**). The cell activated pathways include some signaling pathways

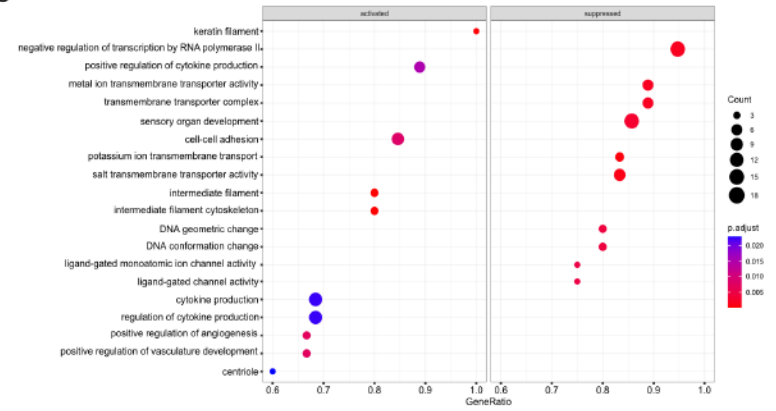
A



B



C



D

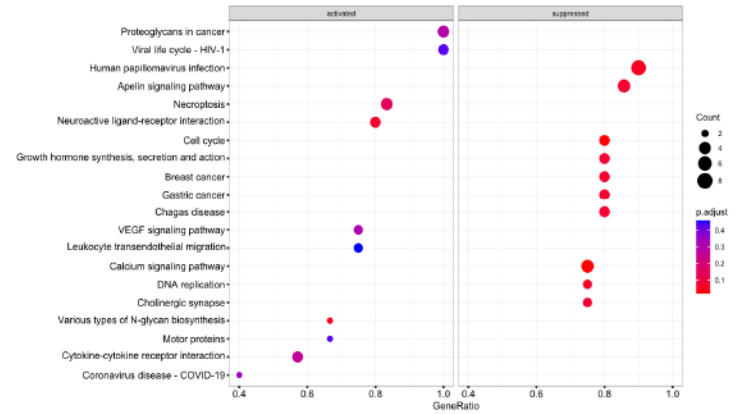


Fig. 5.6. Gene set enrichment analyses (GSEA) shows that cells tend to retain general cellular maintenance and immune evasion gene sets while immune related and anti-proliferation gene sets are packaged in EVs. GSEA dot plots for 1220 cell exclusive RNAs inputted into plots for (A) Gene ontology (GO) and (B) GSEA KEGG pathways; and for 2669 EV exclusive RNA for (C) Gene ontology (GO) (D) GSEA KEGG pathways.

associated with growth and development and associated with tRNA biogenesis. Suppressed pathways included ‘morphine addiction’, ‘GABAergic synapse’, ‘renin secretion’, ‘synaptic vesicle cycle’, ‘cGMP-PKG signaling pathway’, ‘O-glycan biosynthesis’, ‘PD-L1 expression and PD-1 checkpoint pathway in cancer’ and ‘FoxO signaling pathway’ (**Fig. 5.6B**). Thus, we can surmise that cell exclusive pathways are associated with cell maintenance including cell cycle. Additionally, suppression of the PD-L1/PD-1 pathways also show an immune evasion phenotype in the cells.

For EV exclusive transcripts, GSEA analysis revealed increases in GO associated with ‘keratin filament’, ‘cell-cell adhesion’, ‘intermediate filament cytoskeleton’, ‘cytokine production and regulation of production’, ‘positive regulation of angiogenesis’ and ‘centriole’ (**Fig. 5.6B**). Therefore, we can surmise that gene sets associated with cell shape and internal organization, as well as regulating cell growth and activity are packaged into EVs. Conversely, suppressed GO linked to the ‘transmembrane transporter activity’, ‘ligand-gated channels’, ‘sensory organ development’, ‘DNA geometric and conformation change’, and ‘DNA conformation change’ (**Fig. 5.6C**). The suppressed pathways suggests that EV exclusive transcripts = suppress pathways associated with ion flux linked to plasma membrane development and epigenetics. The activated pathways include ‘proteoglycans in cancer’, ‘viral life cycle – HIV-1’, ‘neuroactive ligand-receptor interaction’, ‘VEGF signaling pathway’, ‘leukocyte transendothelial migration’, ‘N-glycan biosynthesis’, ‘cytokine-cytokine receptor interaction’ and ‘coronavirus disease’ (**Fig. 5.6D**). These activated pathways suggest that EV exclusive RNA transcripts activate pathways associated with cancer and immune activation, similar to the Cytoscape network analysis. The suppressed pathways include human papillomavirus infection’, ‘apelin signaling pathway’, ‘cell

cycle', 'growth hormone synthesis, secretion and action', 'chagas disease', 'calcium signaling pathway', 'DNA replication', and 'cholinergic synapse' (**Fig 5. 6D**).

5.2.6 Sequence motifs are shared within cell exclusive and EV exclusive transcripts.

The identified cell exclusive transcripts included mRNAs (75.2%), lncRNAs (10.6%) and ncRNAs (10.2%) (**Fig. 5.7A**). tRNAs made up 1.6%, while miRNAs, snoNAs, rRNA, piRNA, pseudogenes and uncharacterized RNAs made up <1% each. All 1220 exclusive RNAs were ranked by feature count averages and the top 100 genes were extracted and annotated using NCBI database for tRNAs and IGV Refseq sequences for all other biotypes. Using MEME suite, we identified 18 motifs common to the 73 annotated genes from the top 100 genes, which passed the E-value cut-off $E < 0.05$ (**Fig. 5.7B** and **Supp. fig. 5.3**). The width of the motifs that passed the cut-off range between 27 and 50 nucleotides and were shared between 4 to 12 transcripts. Common smaller motifs within the longer motifs included 'AAA', 'GAGA', and 'GGTTC'.

The 2669 EV exclusive RNAs included 79% mRNAs while lncRNAs and ncRNAs made up 10% and 8.9%, respectively. tRNAs made up 1.2%, while miRNAs and uncharacterized RNAs made up 1.4% and 0.3%, respectively. Pseudogenes made up 0.2% while piRNA, snRNA, snoRNAs each contributed 0.04% each to the total number transcripts (**Fig. 5.8A**). As above, the 2669 EV exclusive RNAs were ranked and the top 100 RNAs were annotated. Using MEME suite, we identified 7 motifs that passed the E-value cut-off of < 0.05 that were between 5 and 44 nucleotides (**Fig. 5.8B**, **Supp. fig. 5.3**). The most common motif was present in 71 of the 88 annotated transcripts (82%) and contained the reverse complement of the 'GGAG' packaging sequence, a 'GCCNCC' motif and an overrepresentation of Cs. The second most common was present

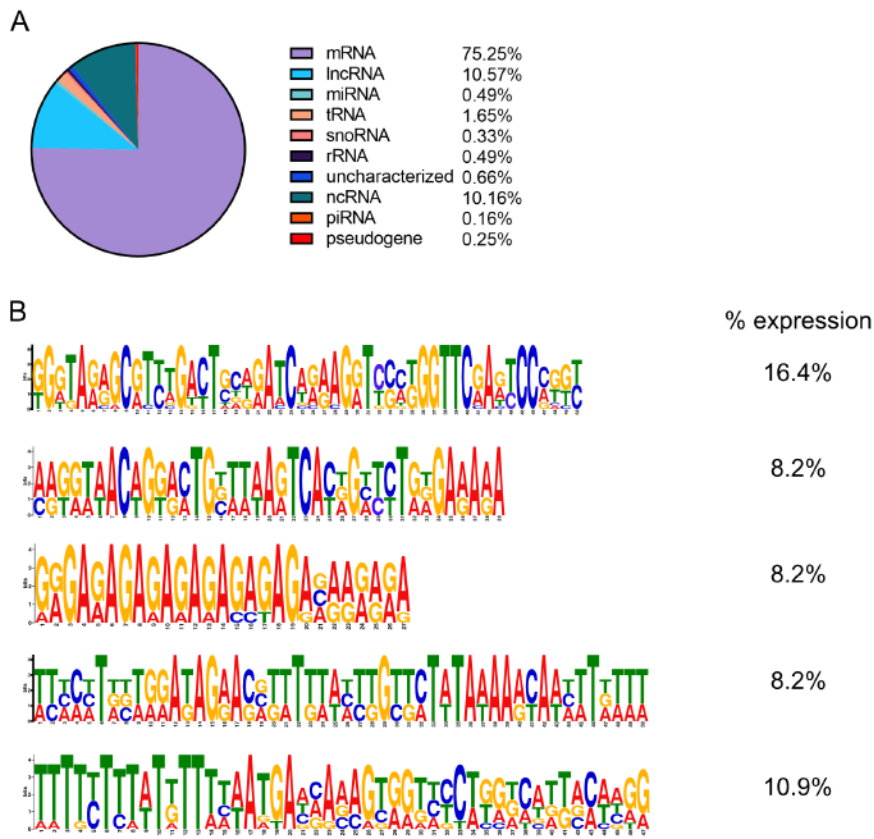


Fig. 5.7. Cell exclusive RNAs share cell inclusionary motifs. (A) Pie chart showing different RNA biotypes for all cell exclusive RNAs. (B) Cell RNA inclusionary motifs common to top 100 cell exclusive RNAs (ranked by average expression) (E value <0.05).

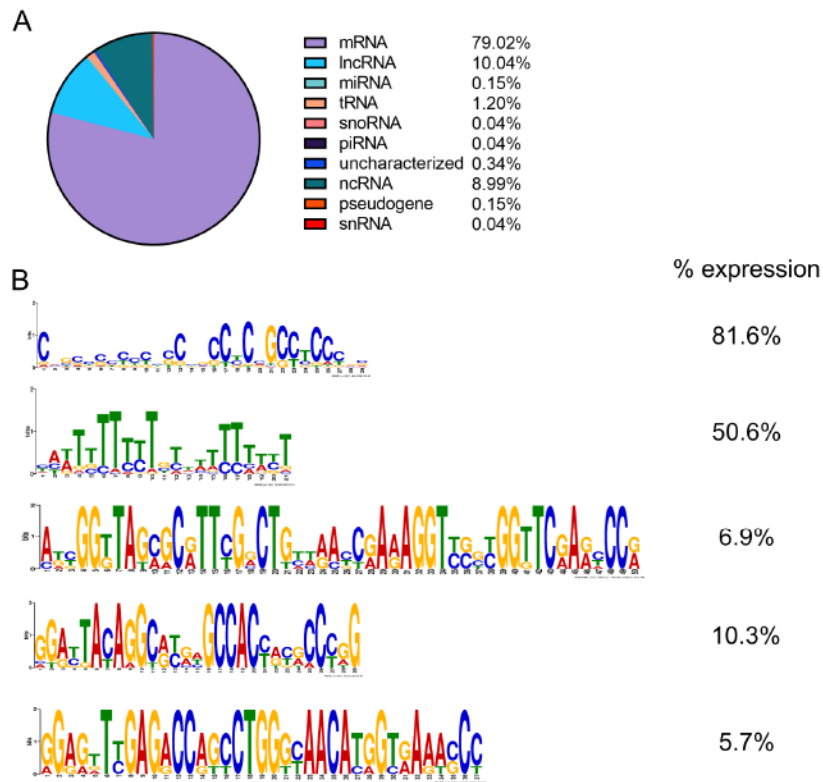


Fig. 5.8. EV exclusive RNAs share EV packaging motifs. (A) Pie chart showing different RNA biotypes for all EV cell exclusive RNAs. (B) EV RNA inclusionary common to top 100 EV enriched RNAs (ranked by average expression) (E value <0.05).

in 44 transcripts (51%) and included a 'UUU' packaging sequence. The packaging sequences from the EVs were generally shorter than the cell motifs.

The two groups of exclusive RNAs (cell and EV exclusive) were individually subtyped into the different RNA biotypes, which were then annotated, and analyzed by MEME suite. For the cell exclusive genes, 5 miRNAs were annotated, and 6 motifs common to these miRNAs were identified (**Supp. fig. 5.3**), none of which met the E-value cut-off of <0.05 . For cell exclusive lncRNA, 75 lncRNAs were annotated and 50 motifs were identified to be common to between 2 - 12 sequences (**Supp. fig. 5.3**), none of which met the E-value cut-off of <0.05 . There were 20 cell exclusive tRNA transcripts with 16 motifs, shared between 2 – 20 RNA transcripts (**Supp. fig. 5.3**), but only 2 motifs passed the E-Value cut-off (**Fig 5.9A**). All 20 shared the most common 50 nucleotide-long motif while 13 shared the second most common 15 nucleotide-long motif. For the EV exclusive tRNA, 30 tRNAs were annotated (**Supp. fig. 5.4**) and five motifs passed the E-value cut-off, shared between 5 – 22 tRNAs with 73% of the tRNAs sharing the most common motif (**Fig 5.9B**). The most common cell-exclusive and EV-exclusive tRNA motifs were nearly identical suggesting that this is a tRNA motif and not an inclusionary/exclusionary motif with respect to EV packaging. For the EV exclusive lncRNAs, 141 lncRNAs were identified (**Supp. fig. 5.4**) and only 2 motifs passed the E-value cut-off with 17% sharing the most common motif, which was the poly-U motif (**Fig 5.9C**).

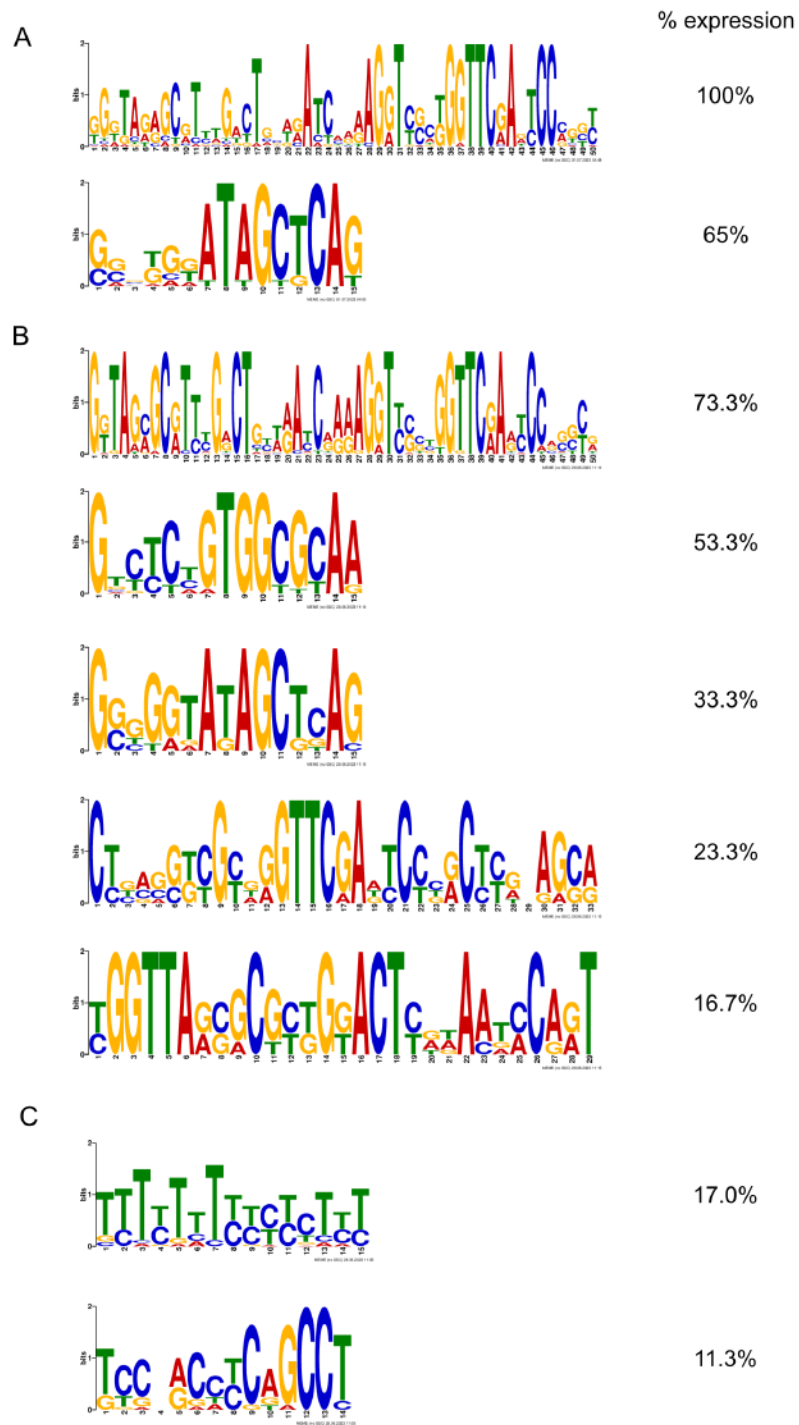


Fig. 5.9. RNA biotype specific motifs for (A) full length cell exclusive tRNA (B) full length EV exclusive tRNA (C) RNaseq extracted fragments for EV exclusive lncRNA. (E value <0.05).

5.2.7 Different lncRNA and tRNA fragment subtypes are packaged into EVs.

Of the different types of lncRNAs packaged from cells to EVs, long intergenic non-coding RNAs (LINC)s make up the highest contribution, followed by antisense transcribed (AS), unclassified lncRNAs, divergent transcripts, miRNA-associated host gene lncRNA (HG), and intronic lncRNAs being the least abundant (**Fig 5.10A, Supp. file 5.2**).

Interestingly, tRNAs were packaged into EVs at a high rate with both essential and non-essential amino acid anti-codon types packaged (**Fig. 5.10B, Supp. file 5.2**). Of the tRNA-anticodon fragments identified, the most prevalent anticodon was for Val (nonpolar), followed by Gln (polar, uncharged), Gly (nonpolar), and Lys (positively charged basic). Interestingly, tRNA for Asp, His, and Trp were not found to be packaged. The tRNA's identified to be packaged from cells to EVs were queried against the tRNAs in the human genomes. Of the 596 tRNAs in the genome, 19.2% (115) were found to be packaged into EVs, while 1 packaged tRNA (Val-CAC-1-7) was not found within the genome list. These tRNA-anticodons were not packaged according to the number of tRNAs found in the genome, therefore, appear to be packaged in a gene specific-manner (**Fig. 5.10B**).

5.3 Discussions

In this study, we found that multiple RNA subtypes are packaged into pediatric B-ALL EVs including miRNAs, lncRNAs, mRNAs, snRNAs and tRNAs. Packaging in this context is via a combination of RNA loading and sorting mechanisms. Furthermore, the selectively packaged (sorted) RNA transcripts have motifs that are shared between many of the RNA sequences.

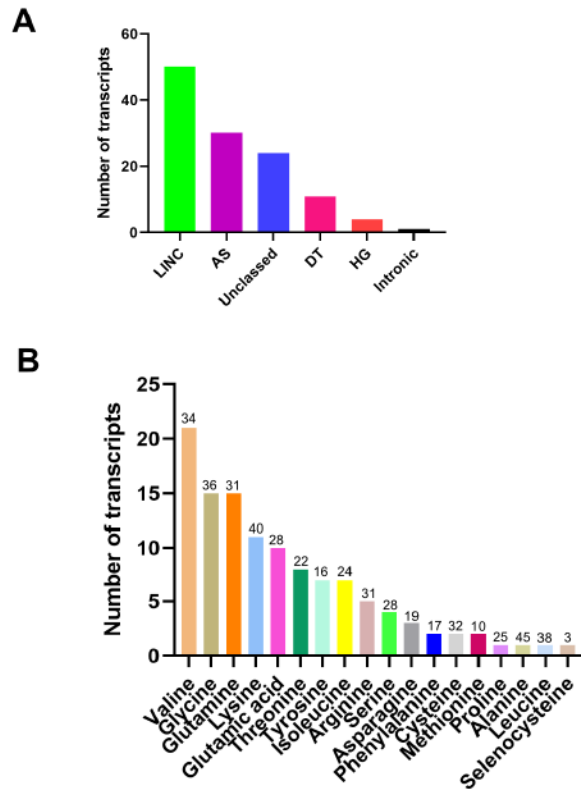


Fig. 5.10. Different types of lncRNA are packaged in EVs and tRNA are differentially packaged depending on the amino acid. (A). The number of packaged transcripts of the different types of lncRNAs that are packaged into EVs are shown. Long intergenic non-coding RNA (LINC), divergent transcripts (DT), antisense transcripts (AT), intronic, microRNA host genes (HG) and unclassified lncRNAs. (B) The number of packaged tRNA-anticodon transcripts are shown as columns, with the total number of tRNA genes in the genome indicated by the number above the column.

Notably, we identified the poly-U motif previously described along with a novel GCCNCC motif in packaged transcripts.

Using sRNADE analysis, read counts of cells and their EVs showed that counts from immortalized cell lines and primary lymphoblasts do not differ significantly. Overall, counts from lymphoblasts were all close to 1×10^7 . All counts from EVs were $> 1 \times 10^7$. This suggests that small RNASeq was able to capture a comprehensive picture of RNA transcripts comparable between cells and EVs. Interestingly, different types of RNA transcripts were packaged into the EVs. The fact that at least 10% of the total RNA (depending on the sample) currently needs to be annotated into broad categories is a noteworthy observation. This suggests that the full interpretation of our data is currently limited by the annotation gap in the field. Therefore, more work needs to be done within the RNA community to investigate and validate RNA transcripts. Also, of all those annotated, at least 10% are noncoding RNAs that are currently unclassified. This is in line with previously published work by other groups using different study models. For instance, recently, O'Grady and co-workers using EVs from HUVEC cells found at least 10% of RNA in cells and EVs are classed as other⁶³². Also in 2020, Zhao and colleagues found $>10\%$ unannotated RNA population in their RNASeq of serum EVs from humans, rats and mice⁶⁶⁰.

Of the annotated and classified transcripts, the most abundant RNA species in cells and EVs were miRNA, mRNAs (sense and antisense), tRNA followed by lncRNAs. These RNA transcripts were both fragments (for larger transcripts like mRNA and lncRNAs) and full-length miRNAs. mRNAs, lncRNAs, and miRNA have previously been reported to be present in EVs from mouse hypothalamic neuronal GT1-7 cell lines⁶⁶¹; serum from male Sprague Dawley rats, male C57BL/6 mice and healthy human volunteers⁶⁶⁰; serum of patients with chronic hepatitis and acute-on-chronic liver failure caused by HBV⁶⁶²; healthy human plasma⁵²²; HEK293T, Neuro2a and

hTERT-MSC cell lines⁵³²; and plasma of HIV-negative and -positive donors⁶⁶³. We found a very low level of ribosomal RNA transcripts present in cells and their EVs. This was similarly shown in a study by Chen and Co-workers where they found rRNA in serum EVs from patients with chronic hepatitis and acute-on-chronic liver failure⁶⁶². Huang and colleagues also found rRNA in EVs from healthy human plasma⁵²², while Quek and colleagues found rRNA in EVs from mouse hypothalamic neuronal GT1-7 cell lines⁶⁶¹. Previously, using a different RNA isolation technique, we have also found rRNA present in EVs³⁵¹.

Our finding that miRNAs are the most abundant small RNA species in EVs from pediatric B-ALL lymphoblast EVs is consistent with other studies^{629,664}. However, a few publications recently showed that the abundance of miRNAs in EVs is much less than previously reported, specifically stating that miRNAs make up only a small proportion of EV RNA cargo in B cell lines and HEK293 cells⁵³³ and from a broad range of cancer cells⁶⁶⁵. Taken together, our data and results from published data clearly show that different RNA biotypes are packaged into EVs with **miRNA** not being the majority.

The sRNADE data, which allowed for analysis of miRNA specifically, showed that both mature miRNA and precursor miRNA were packaged from cells into EVs. Pre-miRNA was previously reported to be packaged into EVs from human androgen-independent prostate cancer PC3 cells⁶⁶⁶; into LOX melanoma cells⁵⁴¹; breast cancer cells⁶⁶⁷; bone marrow-derived macrophages⁶⁶⁸ and mesenchymal stem cell EVs⁶⁶⁹. The packaging of precursor miRNAs begs the question about whether the EV packaged pre-miRNAs are attached to or detached from the RISC complex. It is possible that cells use different mechanisms for loading and sorting pre- and mature miRNAs into EVs. It has been shown in breast cancer cell lines MCF10A and MDA-MB-231 that components of the miRNA biogenesis pathway, such as RISC proteins, AGO2, Dicer and TRBP, were detected

in cancer exosomes but not normal cell exosomes (normoxomes)⁶⁶⁷. Additionally, in LOX melanoma cells, the ARF6-Exportin-5 complex shuttles pre-miRNA into tumour microvesicles⁵⁴¹. This suggests that pre-miRNAs could potentially be packaged into EVs and processed within EVs. Overall, it remains to be determined how pre-miRNA and miRNA are packaged into B-ALL EVs and whether pre-miRNA is processed within the B-ALL EVs.

Non-selective packaging (loading) of RNAs from cells and EVs may be a form of constitutive packaging that is simply concentration dependent. A non-selective mechanism has previously been described in MDA-MD-231 breast cancer cell²⁹⁸ and MCF-7, MCF-10, and MCI-H1299 breast cancer cell lines⁶⁴⁵ and lung carcinoma cell line A549⁶⁷⁰. It is possible that non-selective packaging of RNA transcripts may be a regular process that is part of normal physiology. Our finding that the non-selectively packaged transcripts are associated with genes linked with cellular metabolism supports this hypothesis.

Selective packaging involves an exclusionary or inclusionary mechanism where RNA transcripts are present in cells but excluded from EVs or *vice versa*. This selective RNA packaging mechanism aligns with previously published work in murine and human mast cells⁶²⁸, ovarian cancer⁶⁷¹, antigen-presenting cells⁶⁵⁵, lung adenocarcinoma⁶⁷² and glioblastoma⁶⁷³. Selective packaging (EV or cell exclusive) may be an active process that may be dependent on specific signals. One possible signal is the presence of selectivity motifs within the packaged RNA. These motifs could be the link between the RNA to be packaged, the associating RBPs, and the packaging process. Similar packaging motifs were found in primary mice neurons and astrocyte small EVs which were called core EXO-motifs (UGUG(U/C), CAUG and CNGGNG) and core-cell motifs such as AGAAC, CAGU and AUAA that affected RNA sorting into EVs or retention in cells²⁹⁵. The O'Grady group also identified large motifs, which contain smaller motifs such as the GGAG and CGCGC using

the human umbilical vein endothelial cell line⁶³²⁶³². Finally, miRNAs packaged from primary and immortalized neurons and astrocytes into EVs shared motifs such as GUAC and CACACA⁶⁷⁴. Using MEME suite, we identified >20 nucleotide long motifs which are shared by at least 6/100 (6%) of the cell exclusive RNA transcripts. Furthermore, the ‘AAA’, ‘GAGA’, and ‘GGTTC’ motifs within the large motifs did not appear with the EV motifs. With respect to EV exclusive RNA transcripts, one motif is shared by 81% of the top EV exclusive RNAs, while another motif is present in the top 50%. The second most common motif, which is 21 nucleotides, contain the poly-U motif previously identified to be present in EV-retained miRNA⁶⁵³ while the most common motif contained an overrepresentation of Cs, and the ‘GCCNCC’ motif.

Our data suggests that cells retain RNA transcripts needed for maintaining their physiology and for regular metabolism. Additionally, genes were found that are associated with embryonic heart tube morphogenesis, myeloid cell development, synapse maturation, blastocyst formation, stem cell population maintenance and oocyte maturation suggesting that the cells also retain genes important for cellular differentiation, which may be associated with the developmental arrest experienced in B-ALL.

While some signalling pathways were identified in the cell exclusive transcripts, there were more signaling pathways associated with the EV exclusive transcripts. Also, the signaling pathways in the cell exclusive transcripts were associated with cell development and homeostasis, while the signaling pathways in the EV exclusive transcripts are largely associated with disease, including B cell leukemia. Some pathways that are present in both cell exclusive and EV exclusive transcripts, include gene expression (transcription and translation) associated pathways. For instance, both cells and EVs carried RNA associated with cell division (e.g. nucleosome), which

may suggest that some cell division associated genes may be retained while others will be shuttled out into the EVs.

EV exclusive transcripts include disease relevant genes such as genes for leukocyte trans endothelial migration, TGF β signaling pathway, MLL1/2 complex, interferon gamma pathway, IL-6 production, and $\alpha\beta$ T cell production. Additionally, genes associated with gene expression and those indicative of infections such as Yesinia and Shigellosis are identified in the pathway analyses. Furthermore, a notable pathway in the EV exclusive cohort is the '3'-untranslated region (3'UTR) AU-rich region binding'. The alteration of the mRNA 3'UTR AU rich region that contain AU-rich elements (AREs), is found to be linked to mRNA transcript instability, which is associated with cancer and disease progression⁶⁷⁵⁻⁶⁷⁷. Specifically, mRNA 3'UTR AU rich binding is linked to pediatric B-ALL. For example, PAX5, a transcription factor that regulates B cell development, has been shown to have prevalent 3'UTR editing events in healthy B lymphocytes⁶⁷⁸. However, in B cell cancers, the PAX5 3'UTR is shortened, leading to an increase in oncogenic translation frequency, which is in turn associated with disease progression⁶⁷⁸. Similarly, transendothelial migration is increased in B-ALL, especially in disease and bone marrow relapse⁶⁷⁹. Furthermore, transendothelial migration, in addition to cell adhesion is associated with lymphoblasts infiltration into other organs, infiltration of extramedullary tissues, and linked to relapse⁶⁷⁹. Therefore, EVs in this context may be functioning as vehicles for transporting disease-associated RNA to other B cells or stromal cells to promote bystander effects and disease progression.

We identified different types of lncRNAs packaged into pediatric B-ALL EVs. LINC are the most packaged lncRNAs followed by antisense transcripts, unclassified lncRNAs, divergent transcript, whereas miRNA host genes and intronic lncRNA are the least packaged. The high prevalence of LINC suggests that they may be playing specific roles. LINC RNAs exist within the introns of

coding genes, which means they can disrupt gene function by sequestering the target genes^{593,680}. Some studies had previously identified lncRNAs as important regulators of leukemogenesis in B-ALL, including regulating RNA:RNA interactions⁶⁸¹⁻⁶⁸⁶. This detection of lncRNA fragments, especially of lncRNAs such as ADORA2A-AS1 in chronic myelogenous leukemia⁶⁸⁷ and BALR-6 in B-ALL⁶⁸⁸, may suggest a mechanism for cells to advance leukemia pathology via EVs.

Finding tRNA anticodon fragments was unexpected. This novel finding raises more questions about the types of RNAs packaged into EVs, and how this may be driving disease onset and progression. Liu and coworkers found alterations in concentrations of amino acids Val, Thr, and Glu in the serum of pediatric B-ALL patients as opposed to controls⁶⁸⁹, all of which we found to be highly packaged in B-ALL EVs as tRNA anticodons. This may be significant because it has been shown that amino acid stress response genes lead to the release of enzymes that may alter sensitivity to drugs such as L-asparaginase⁶⁹⁰. In diseases such as breast cancer, tRNA fragments have recently emerged as a predictive biomarkers for Trastuzumab drug resistance⁶⁹¹. Therefore, pediatric B-ALL EV-associated tRNAs may represent a new source of biomarkers for monitoring disease onset and progression.

5.4 Conclusion

The analysis of the small RNA cargo of pediatric B-ALL primary lymphoblasts and immortalized cell lines is important for understanding and exploring the potential of EVs. The literature led to the hypothesis that mostly miRNAs will be packaged into the EVs in our study. However, small RNASeq showed that in addition to mature miRNAs, precursor miRNAs and fragments of mRNA, lncRNA, tRNAs, snRNAs and snoRNAs and full-length transcripts of some genes were abundantly packaged into EVs. Packaging in this context includes non-selective loading and

selective sorting, as previously identified for other cell types. Furthermore, GO and pathway analysis suggests that EVs may be involved packaging and dissemination of disease specific RNAs. We also identified motifs shared by cell exclusive or EV exclusive transcripts, which may help to engineer EVs that carry-out specific functions upon delivery to recipient cells.

Chapter 6: The effect of Streck preservative treatment of blood collection tubes on EV collection

6.1 Introduction

Of all the potential biofluid biomarkers as candidates for liquid biopsy, EVs are very promising, in many ways more than the others. For instance, because tumors secrete EVs rapidly, tumor biofluids contain millions of circulating EVs⁵³⁴. Furthermore, tumor EVs carry specific bioactive cargo of DNA, mRNA, ncRNA and proteins, suggesting that tumor EVs may offer insights into tumor biogenesis, useful for disease monitoring, prognosis and therapeutics³⁰⁰. Therefore, sampling biofluids for EV analysis is emerging as the next generation liquid biopsy approach.

Blood plasma and urine are two readily available biofluids, that can be obtained through minimally invasive sampling methods. However, because the EV field is emergent, not much is known about the exact parameters for blood and urine collection that is most feasible for EV analysis. Consequently, the effect of pre-analytical processing on biofluid EVs is not well studied or a standard established. For EV collection, pre-analytical processes include the draw/collection process, post collection handling, collection site to processing site transport, post draw/collection processing (centrifugation to deplete cells, or the lack thereof), post processing storage (long- or short-term) and finally EV isolation. These steps are all potential variables that will need to be standardized, before EVs can be used as the biomarker of interest for liquid biopsy.

One key variable, which is the first in the series of pre-analytical processing is sample collection. For blood collection for diagnostics, tubes used are color-coded, based on the type of anticoagulant coating the tube (and the manufacturer), and the type of downstream analysis⁶⁹². For blood collection with the goal of collecting plasma, the anticoagulants of choice are sodium citrate and

ethylene diamine tetra acetic acid (EDTA)^{693,694}. For plasma glucose and lactic acid testing, potassium oxalate is used as the anticoagulant, in addition to sodium fluoride, an antiglycolytic agent⁶⁹⁵. There are also tubes where no anticoagulant is used, which allows the blood to clot.

For urine collection, the main guidelines are to prevent or reduce contamination by organisms from the urethra, skin, genitalia and/or fecal flora, by storing specimens at temperatures <15°C, with a 2 h window from collection to processing⁶⁹⁶. While the effect of these collection tubes have been tested for other liquid biopsy biomarkers, not much is known about their effect or lack of for EVs.

To facilitate collection of blood plasma and urine, a new collection tube called Streck cell-free BCT® (Streck, USA) was developed for the collection of cell free DNA. The basis of Streck BCT is a proprietary stabilization buffer, which stabilizes biofluids for up to 14 days post collection between 6 °C – 37 °C⁶⁹⁷. To date, this tube has not been tested on EVs and their cargo. Therefore, the objective of this study was to characterize blood and urine EVs collected in non-Streck versus Streck BCT; to ascertain if there were differences in their EV amounts, cargo, and contaminants.

6.2 Results

6.2.1 Samples collected in EDTA and Streck cell-free BCT.

Two sets of samples were collected from patients: blood plasma and urine. For blood, samples for each donor were collected in EDTA-coated or Streck BCT containers, respectively. For blood plasma, three donors submitted their samples, while for urine, seven donors submitted. The donor sample details are shown in **Table 6.1** for plasma and **Table 6.2** for urine.

Table 6.1. Details of plasma samples used for study

Sample set	Patient ID	Gender	Clinical Diagnosis	Collection container	Volume (mL)
Set 1	BBN-189	Female	Osteosarcoma	Ethylenediaminetetraacetic acid (EDTA) Vacutainer	0.5 (x4)
Set 1	BBN-189	Female	Osteosarcoma	Streck blood collection tube (BCT)	0.5 (x4)
Set 2	BBN-191	Male	Osteosarcoma	EDTA Vacutainer	0.6 (x4)
Set 2	BBN-191	Male	Osteosarcoma	Streck BCT	0.5 (x4)
Set 3	BBN-198	NA	Spindle cell sarcoma	EDTA Vacutainer	0.6 (x4)
Set 3	BBN-198	NA	Spindle cell sarcoma	Streck BCT	0.5 (x4)

Table 6.2. Details of urine samples used for study.

Sample set	Patient ID	Age (yrs)	Clinical Diagnosis	Collection container	Tumour details
Set 1	SM-5062	57	Poorly differentiated carcinoma	Non-Streck (NS)	metastasis
Set 1	SM-5062	57	Poorly differentiated carcinoma	Streck treated tube	metastasis
Set 2	SM-5070	64	Lung adenocarcinoma	NS treated tube	metastasis
Set 2	SM-5070	64	Lung adenocarcinoma	NS treated tube	metastasis
Set 3	SM-5063	70	Glioblastoma IDH-wildtype	Non-Streck treated tube	primary
Set 3	SM-5063	70	Glioblastoma IDH-wildtype	Streck treated tube	primary
Set 4	SM-5065	3	Rhabdomyosarcoma	NS treated tube	recurrent
Set 4	SM-5065	3	Rhabdomyosarcoma	Streck treated tube	recurrent
Set 5	SM-5072	9	Osteosarcoma	NS treated tube	metastasis
Set 5	SM-5072	9	Osteosarcoma	Streck treated tube	metastasis
Set 6	SM-5067	56	Glioblastoma	NS treated tube	recurrent

			IDH-wildtype		
Set 6	SM-5067	56	Glioblastoma IDH-wildtype	Streck treated tube	Recurrent
Set 7	NA	NA	Healthy volunteer	NS treated tube	NA
Set 7	NA	NA	Healthy volunteer	Streck treated tube	NA

6.2.2 Comparison of blood plasma and urine EV morphology between collection tubes

Using transmission electron microscopy (TEM), the morphology of EVs in urine collected in non-Streck versus Streck BCT is shown (**Fig. 6.1**). The shape and the size of the EVs is similar between non-Streck and Streck tube urine.

Using Nanoparticle tracking analysis (NTA), EV diameter and concentration were quantified in plasma and urine respectively. For blood plasma, the mode and mean diameter did not change significantly between both collection tubes (**Fig. 6.2A, B**). However, the concentration did change, being higher in EDTA tubes than in Streck BCT (**Fig. 6.2C**), with this difference being statistically significant. The same trend (higher concentration in EDTA tube collected than in Streck BCT) was seen for log transformed concentration between plasma collected using both containers (**Fig. 6.2D**), with the difference between both collection tubes being statistically significant.

In urine, for differences in mode diameter between non-Streck and Streck collected urine (**Fig. 6.3A, B, C**), 4/7 samples having a higher mode diameter in non-Streck than Streck and 3/7 having a higher mode diameter in non-Streck than in Streck coated tubes. This trend in the mode diameter was also seen with the mean diameter and concentration. Log transformed concentration for urine collected in both tubes showed that the differences were not statistically significant (**Fig. 6.3D**).

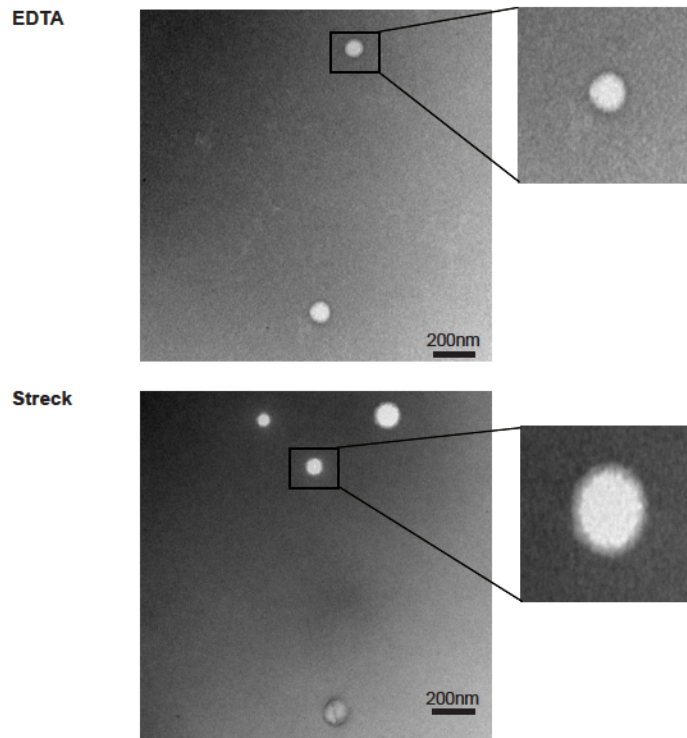


Fig. 6.1. Visualization of urine EVs collected in both collection tubes. The top row is the EVs from the EDTA tube, while the bottom row is Streck BCT collected EVs.

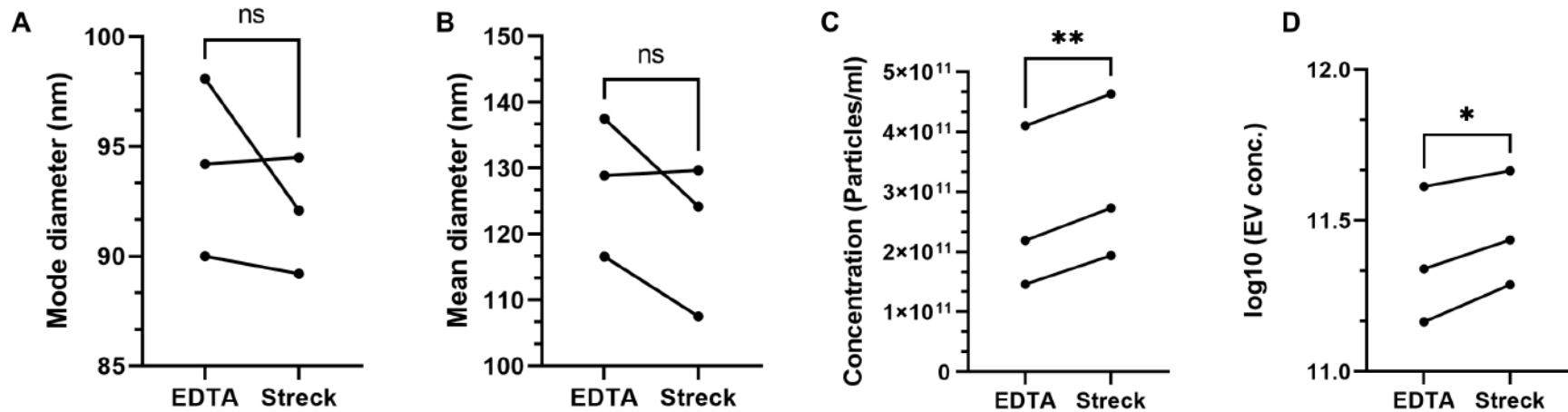


Fig. 6.2. Quantification of EVs in blood samples collected in both sample collection conditions. The samples are shown on the X axis for each histogram, while the diameter (A and B - nm) or the concentration (C – particles/ml) is shown on the Y axis. For each patient, the EDTA versus Streck tube samples were compared for A. mode diameter B. mean diameter C. concentration. D. Log transformed concentration. (Stats = paired t test, p-value <0.05. Difference not statistically significant for the mode and mean, but significant for the concentration and log transformed concentration).

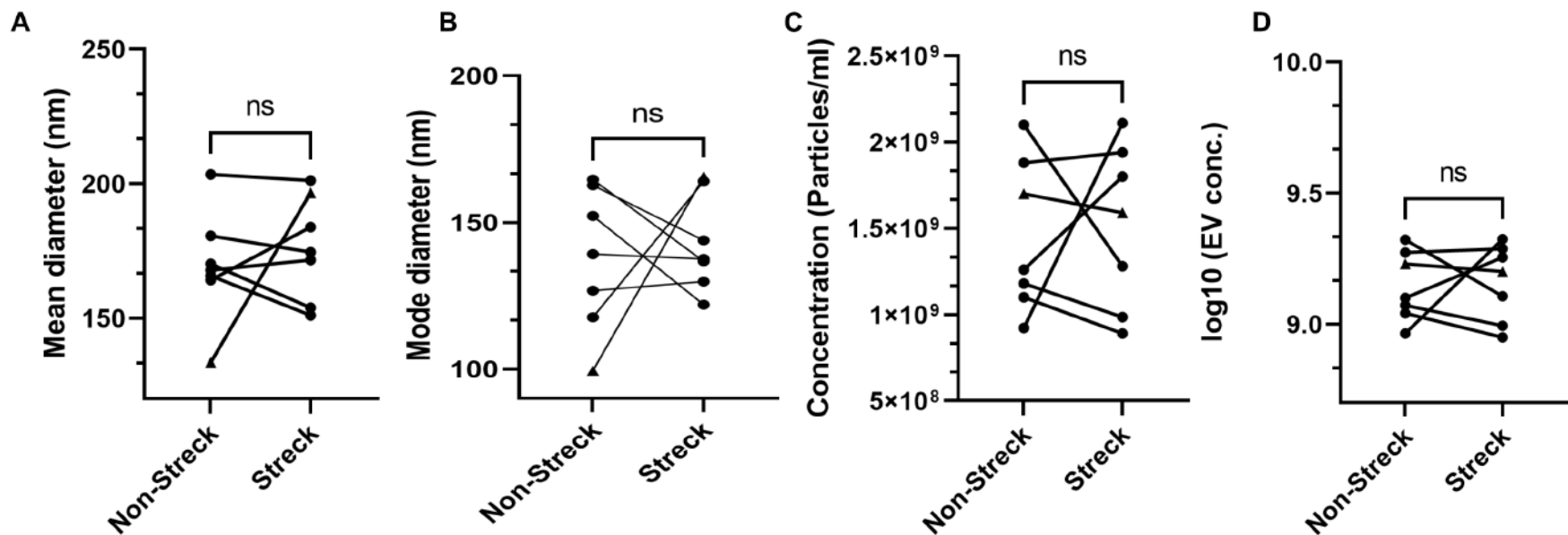


Fig. 6.3. Quantification of EVs in urine samples collected in non-Streck versus Streck coated tubes. For each scatter plot, the patient type is shown on the X axis, while the diameter (A and B - nm) or the concentration (C – particles/ml and log10 (EV concentration)) is shown on the Y axis. (Stats = paired t test, p-value <0.05. Difference not statistically significant). Black circles are disease samples, while triangles are healthy samples.

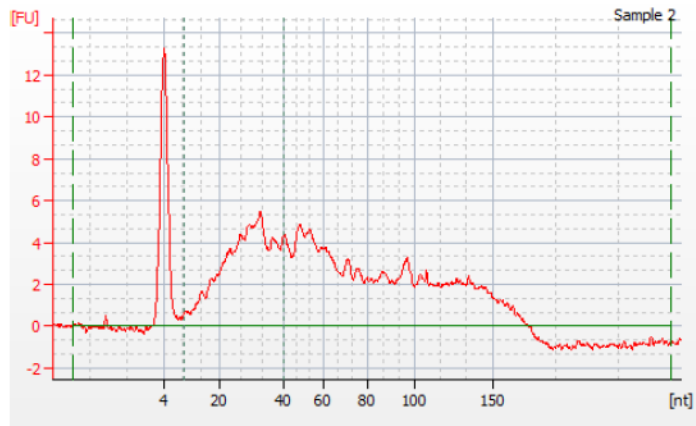


Fig. 6.4. Quantification of plasma EV small RNA. Sample Bioanalyzer output for plasma EV small RNA, showing an electropherogram of EDTA container collected plasma EV small RNA.

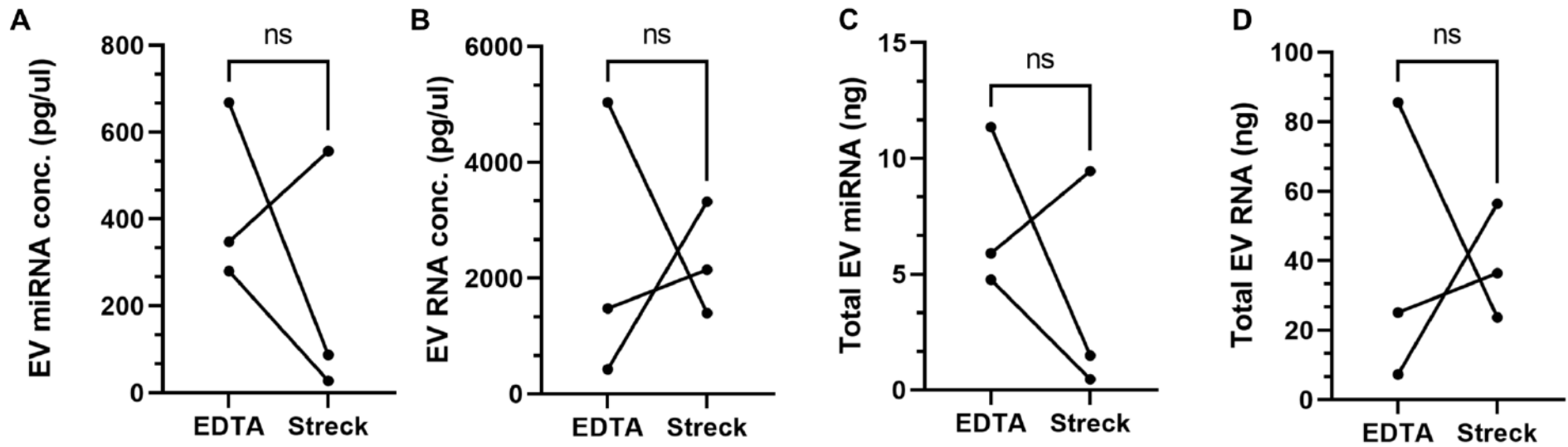


Fig. 6.5. Comparison of the RNA in plasma EVs for (n=3). EV A. miRNA concentration (pg/ul) B. total small RNA concentration (pg/ul) C. total miRNA amount (ng) D. total small RNA amount (ng). (Stats = paired t test, p-value <0.05. Difference not statistically significant).

6.2.3 *Comparison of differences in EV RNA cargo between collection tubes*

For both blood plasma and urine, EV RNA quantification using Agilent facilitated the determination of RNA quantity, useful for downstream steps. A sample bioanalyzer output showing blood plasma EV small RNA output is shown (**Fig. 6.4**).

For blood plasma, the concentration of plasma miRNA (**Fig. 6.5A**), and total RNA (**Fig 6.5B**) were shown to be higher in EDTA-coated tubes than the Streck BCT. This trend is consistent for the total miRNA (ng) of plasma (**Fig. 6.5C**) and the blood plasma EVs (**Fig. 6.5D**), with the difference not being statistically significant.

For urine, the miRNA (**Fig. 6.6A**) and small RNA (**Fig. 6.6B**) concentration and the total miRNA (**Fig. 6.6C**) and small RNA (**Fig. 6.6D**) trends towards being higher in non-Streck versus Streck treated tubes, with the difference between tubes not being statistically significant except for miRNA concentration (**Fig. 6.6A**).

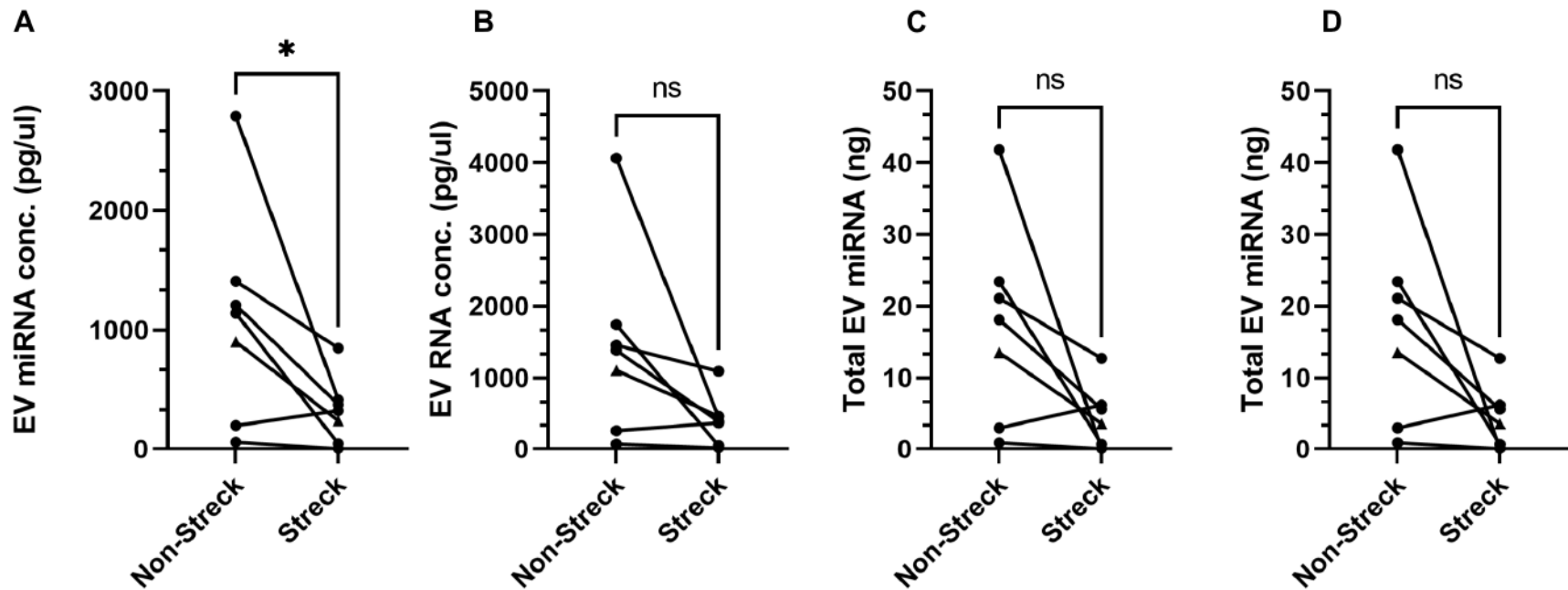


Fig. 6.6. Comparison of the RNA in urine EVs for (n=7). EV A. miRNA concentration (pg/ul) B. total small RNA concentration (pg/ul) C. total miRNA amount (ng) D. total small RNA amount (ng). (Stats = paired t test, p-value <0.05. Difference not statistically significant)

Black circles are disease samples, while triangles are healthy samples.

6.2.4 Comparison of specific EV RNA content

To determine if EVs from both collection tubes have similar non-vesicular RNAs, we did a search for known potential EV contaminating miRNA. Non-vesicular miRNA in this context are miRNAs not associated with the EVs of interest but derived from other surrounding cells or tissues. Furthermore, for accurate interpretation of liquid biopsy EV RNA data, the specificity of the EVs (from cancer cells versus non-cancer cells) can confound the data. For instance for blood, as the liquid biopsy based studies advance in recent years, the need to avoid contamination of blood EVs with EVs from coagulation, platelets or lysed blood cells has led to a preference for plasma⁶⁹⁸. Studies have shown that the potential contamination of plasma EVs with non-vesicular miRNAs is a concern, as it could impact the direction and interpretation of biomarker studies⁶⁹⁹. Also, the potential of contamination by non-vesicular molecules including miRNAs is also a possibility with urine⁷⁰⁰. Pre-analytical variables including biofluid collection contribute to the introduction of non-vesicular contaminants into EV prep^{701,702}. Therefore, it is important to ascertain the effect of Streck versus non-Streck collection tubes on EV isolation from blood plasma and urine.

For plasma, we identified miRNAs known to be involved in platelet activation (miR-126 and miR-223), and a potential normalizer miRNA (let-7b)^{703,704}. Next, we designed primers to detect these miRNAs in EVs from plasma and validated these primers (**Table 6.3, Fig 6.7**); with some primers validating better than others. Finally, we ran RT-qPCR reactions for the one plasma pair that had enough total small RNA with the other two being too low in the Streck tube. Our RT-qPCR data showed that miR-126 and miR-223 levels are higher in EDTA-treated than Streck BCT samples (**Table 6.4**). Furthermore, the normalizer miRNA levels appeared to be at a relatively higher level in Streck BCT than EDTA tubes.

Table 6.3. Primer efficiencies and R² for plasma EV miRNAs.

	miRNA	R²	Efficiency (%)
1	hsa-let-7b	0.996	97
2	hsa-miR-126	0.906	36
3	has-miR-223	0.9701	16

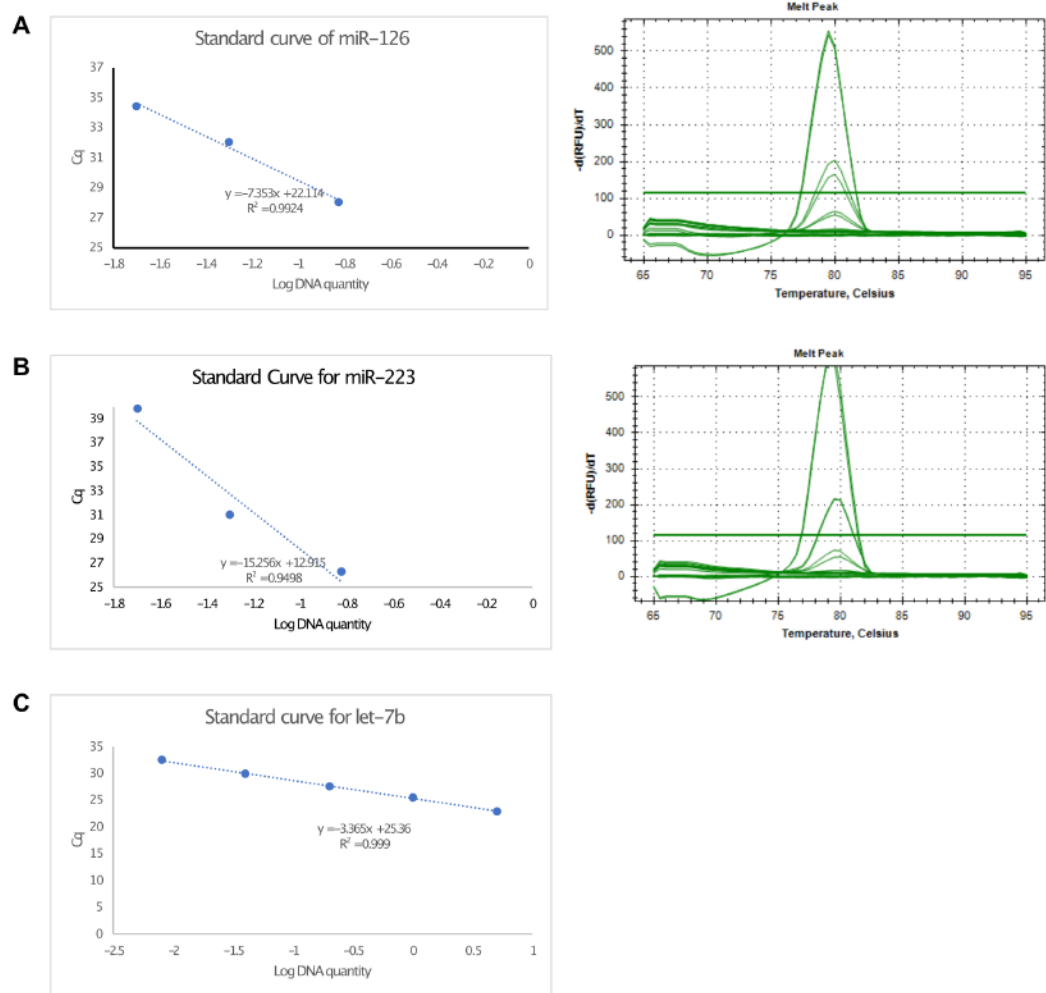


Fig. 6.7. qPCR primer validation for plasma miRNAs. Standard curve, melt curve and melt peak for A. miR-126, B. miR-223 and C. Let-7b.

Table 6.4. qRT-PCR Ct values for three miRNAs from the one plasma sample pair.

	EDTA	Streck
miR-126	28.28	34.72
miR-223	20.83	27.07
Let7b	21.31	18.10

For urine, we identified miRNAs known to be packaged into urine EVs, selected based on publications (**Table 6.5**). We also designed primers to detect these miRNAs in EVs from urine and validated the primers some of which validated as shown in **Table 6.6, Fig. 6.8**. Finally, we ran qRT-PCR reactions for all miRNAs except miR-10b because of the low R2 value during the primer validation, for five urine sample pairs which had enough total small RNA for cDNA synthesis. Our RT-qPCR data showed that all tested miRNAs are higher in non-Streck treated than in Streck treated tube collected urine (**Fig. 6.9**).

6.3 Discussion

For plasma, between the EDTA coated and the Streck BCT tubes, the mean and mode diameter of the EVs were relatively similar. However, for the concentration, the EV contents were higher in EDTA tubes than in Streck BCT. This suggests that the Streck BCT preservative does not adversely affect EV size but seems to have a small statistically significant effect on EV concentration. Thus, we can infer that though Streck BCT and EDTA tubes have different preanalytical effects on plasma and urine EVs, the effects in both cases may be comparable.

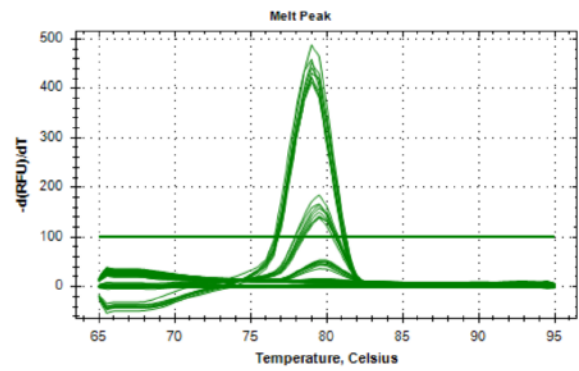
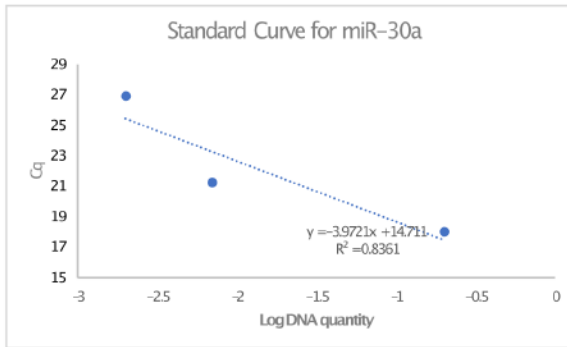
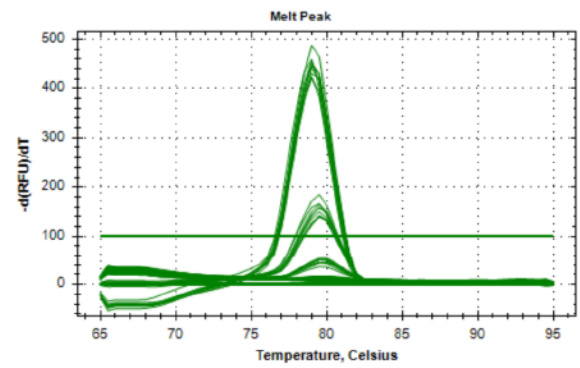
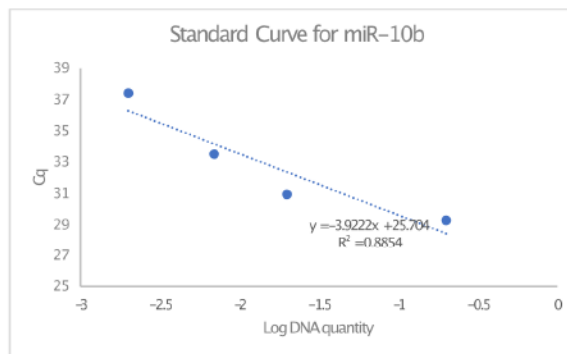
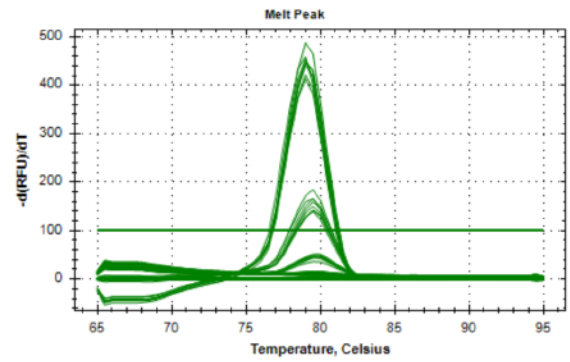
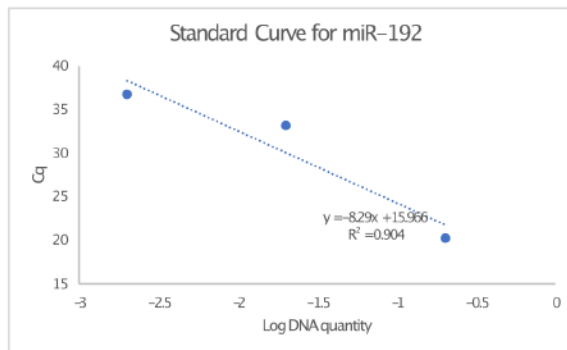
For urine, we determined that the EV mean and mode diameter did not change significantly between collection tubes. These diameters were variable, being higher in non-Streck containers and vice versa for other donor pairs. For the concentration, the concentration is higher in non-Streck tubes than Streck tubes in 3/7 samples, the difference in either direction is not statistically significant. This

Table 6.5. Top 5 selected miRNAs for urine EV qPCR.

	miRNA	Function	Reference
1	miR-192-5p	Abundant miRNA cell-free in urine. Involved in suppression, initiation, and progression of osteosarcoma by targeting USP1	Zhou et al, 2018 ⁷⁰⁵
2	miR-204-5p	Urinary extracellular vesicles miRNA – A new era of prostate cancer biomarkers	Jain et al, 2023 ⁷⁰⁶
3	miR-181a-5p	Promotes Osteosarcoma Progression via PTEN/AKT Pathway. Improves Myogenic Commitment in Murine Fusion-Negative Rhabdomyosarcoma	Sun et al, 2022 ⁷⁰⁷
4	miR-30a-3p	Suppresses the Growth and Development of Lung Adenocarcinoma Cells Through Modulating GOLM1/JAK-STAT Signaling. Correlated with progression and immune infiltration and a potential therapeutic biomarker for lung Adenocarcinoma	Jiang et al 2022 ⁷⁰⁸
5	miR-10b-5p	Plays a pro-tumorigenic role in osteosarcoma cancer	Sheedy & Medarova, 2018 ⁷⁰⁹

Table 6.6. Primer efficiencies and R² for urine EV miRNAs.

	miRNA	R²	Efficiency (%)
1	hsa-miR-181a	0.802	58
2	hsa-miR-30a	0.8361	79
3	hsa-miR-204	0.9798	23
4	hsa-miR-10b	0.8854	80
5	hsa-miR-192	0.904	32

A**B****C**

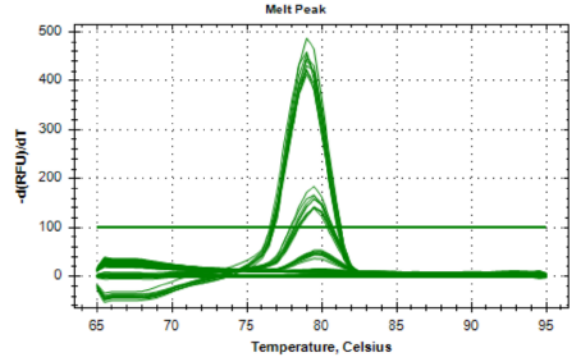
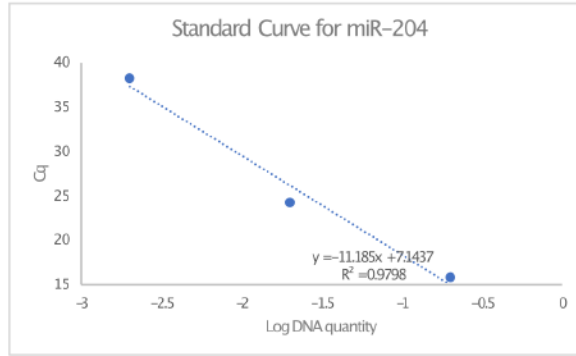
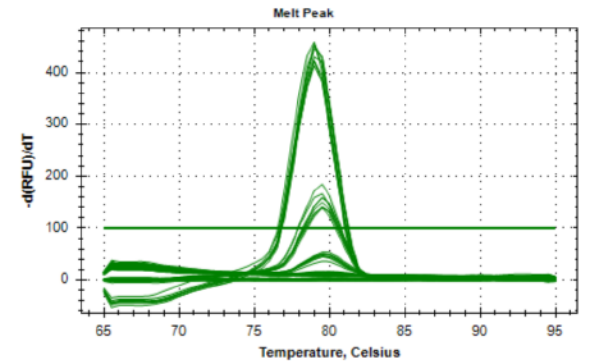
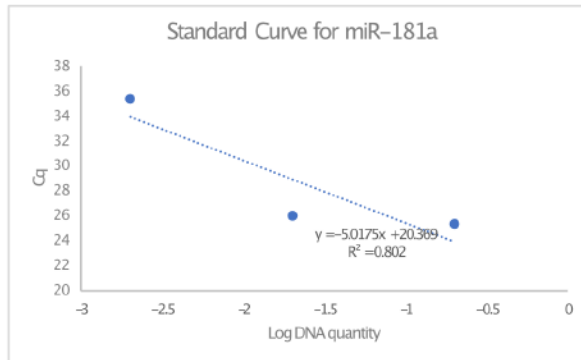
D**E**

Fig. 6.8. qPCR primer validation for urine miRNAs. Standard curve and melt peak for A. miR-30a, B. miR-10b, C. miR-192, D. miR-204, and E. miR-181a.

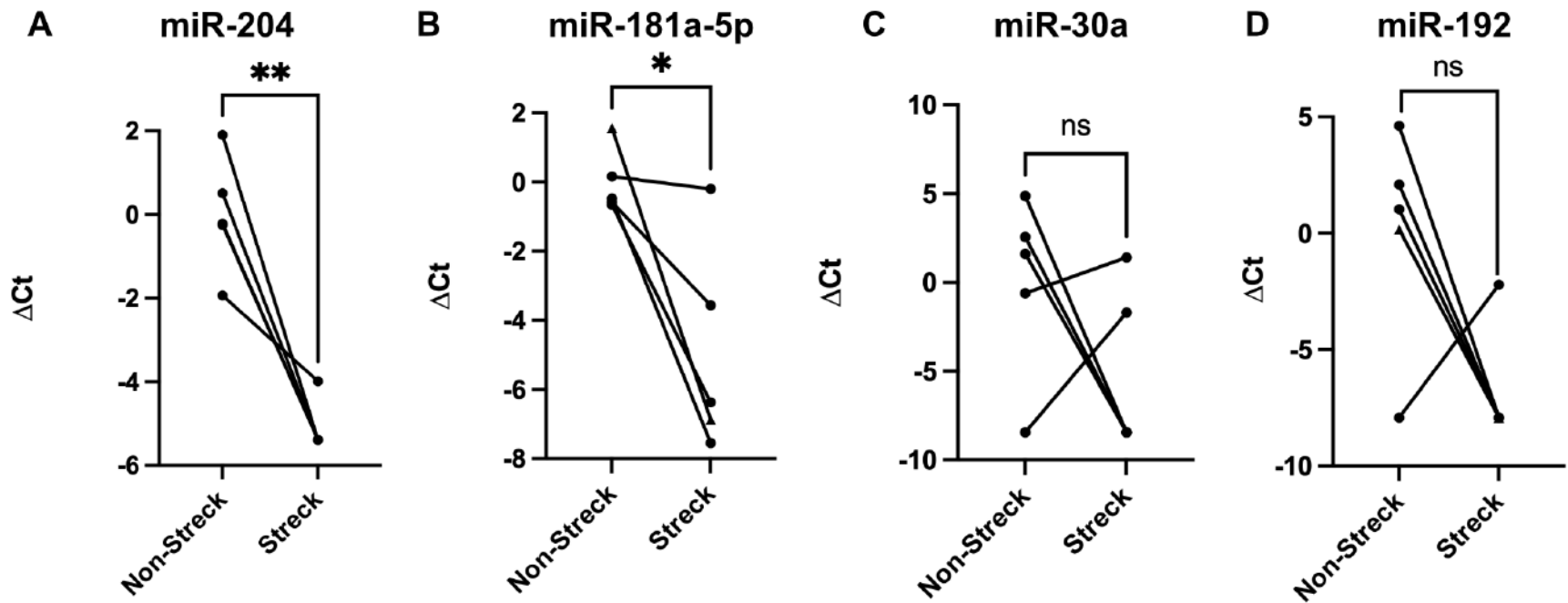


Fig. 6.9. Relative levels of urine miRNAs. Scatter plots of the delta cycle threshold (Ct) of the miRNAs in urine for A. miR-204 B. miR-181a-5p (C. miR-30a (D. and miR-192 (Stats =paired t test, p-value <0.05. Difference not statistically significant).

This data suggests that Streck treatment of urine collection tubes does not have clear adverse effects on urine EV analysis. Furthermore, a comparison of urine EV shape between non-Streck and Streck tubes showed that Streck treatment does not alter urine EV shape. Taken together, Streck preservative does not exert any additional pre-analytical effects on urine EVs.

For EV isolation, we used the synthetic peptide Vn96 which is now well established in our lab and others for isolation of EVs from blood plasma, where it is known that volumes less than 1ml can yield enough EVs for some downstream analysis. However, for urine, the minimum urine volume that would yield enough EVs for downstream analysis is not known. Using the RNA cargo as our parameter, we compared EV RNA amount from 1.25ml, 2.5ml and 5ml of urine (data not shown) and determined that volumes as small as 2.5ml urine can yield an RNA amount that can be used for downstream analyses such as qPCR. Thus, for this study, we isolated EVs from 4ml urine.

In preparation for our RT-qPCR, we attempted to synthesize cDNA from the total small RNA for plasma donors (n=3, 2 sets of collection tubes) and urine donors (n=7, 2 sets of collection tubes). With the starting total small RNA cut-off of 2.3ng, we found that for plasma we could synthesize cDNA from n=1 donor and n=5 donors for urine. In all samples, the EDTA tube collected EVs for urine and plasma were at least 2x higher than in Streck tubes. Based on this information, we recommend that for plasma, Streck BCT collection would need at least 2ml, to yield enough EVs with RNA enough for RNA analysis. For urine, Streck treated tubes would need a minimum of 6ml of urine to yield enough EVs for RNA downstream analysis.

For the quantification of specific miRNAs, the miRNAs we focused on are known to be associated with plasma and urine. For plasma, two markers of platelet activation (miR-126 and miR-223) which are some of the highest in platelets were chosen^{710,711}. The RT-qPCR data showed that miR-126 and miR-223 are higher in EDTA than Streck BCT tubes. However, the plasma housekeeping

miRNA let-7b was higher in Streck BCT than EDTA tubes. This suggests that there is a higher platelet activation in EDTA collection tubes allowing release of EVs by activated platelets post blood collection. However, the Streck preservative seems to reduce or ameliorate this post collection platelet activation and EV release, which is an advantage of Streck preservative over EDTA anti-coagulant. For urine, we focused on miRNAs which may be associated with urinary tract associated cells. Overall, the data showed that miR-204, miR-181a, miR-30a and miR-192 are higher in EDTA treated tube urine than in Streck treated tube urine.

6.4 Conclusion

Overall, Streck preservative seems safe for preserving plasma and urine EVs. However, it reduces the amount of RNA cargo in the EVs. This may be because of interactions between the preservative, the quenching agent and potentially the formaldehyde released and the RNA or proteins on the surface of the vesicles, which may affect the EV luminal RNA content. Aldehydes such as formaldehyde are able to react with a nucleobase to cause nucleic acid denaturation⁷¹². However, because the exact composition of the Streck buffer is not known, we are limited in our interpretation of the meaning of our data. Thus, we recommend that for the purpose of EV analysis, blood collection with Streck tubes requires double the amount needed for EDTA tube collection. This is also applicable in urine collection. While using non-Streck tube has no adverse effect on EV isolation, the potential for contamination with non-vesicular RNA is increased when non-Streck tubes are used for collection. Therefore, we recommend maintaining the current status quo with sample collection with well researched and validated tubes. Although Streck tubes could slowly be incorporated into use, more research is needed to answer questions such as whether the degradation of RNA due to Streck treatment is uniform, or not.

Chapter 7: Discussion, next directions, and significance

7.1 Discussion

The overall objective of this study was to explore the role of miRNAs and extracellular vesicles in specific human diseases (such as pediatric B-ALL and carcinoma). To achieve this objective, we investigated miRNAs in cells as well as in EVs associated with two human biofluids, peripheral blood and urine. Firstly, we analyzed the literature to identify miRNAs that were discriminatory for pediatric B-ALL. However, we are not able to find a clear discriminatory signature that might be present in cells and, therefore, in EVs. Next, we isolated and characterized blood plasma derived EVs in pediatric B-ALL, with the aim of finding RNA profile that discriminates pediatric B-ALL from non-disease controls. Next, we profiled the EV RNA cargo of primary pediatric B-ALL lymphoblasts and immortalized pediatric B-ALL cell lines, with the aim of exploring RNA packaging from B-ALL cells to EVs. The last aim was to investigate the effect of pre-analytical processing on EV characteristics. Specifically, we aimed to compare the effect of collection tube pre-treatment with Streck preservative versus EDTA anticoagulant on EV characteristics.

To identify miRNAs that were discriminatory for pediatric B-ALL (differentially expressed miRNAs (DEmiRs)), we performed a reanalysis of five publicly available miRNA datasets followed by a *meta*-analysis of previously identified DEmiRs from 25 studies. Overall, the reanalysis showed that the DEmiR data clustered by platform and not by disease state. The *meta*-analysis also did not reveal a consensus miRNA signature as there were many miRNAs upregulated in some studies and downregulated in others. However, eight promising miRNAs (miR-181b, miR-128b, miR-181a, miR-128, miR-128a, miR-181c, miR-155, miR-142-3p, and miR-451) were identified from the *meta*-analysis, which could be the basis of future investigations.

This study revealed that standardization of miRNA isolation and analysis is needed in B-ALL to enable cross-study comparisons and identification of a consensus signature.

To identify an RNA profile that discriminates pediatric B-ALL from non-disease controls, we used pediatric B-ALL and NCD plasma. Using NTA, we found that B-ALL plasma contains more EVs than NCD plasma. Then, using a synthetic peptide-based isolation technique (Vn96), which is clinically amenable and able to isolate a comprehensive EV population, we isolated EVs from NCD and pediatric B-ALL blood plasma, followed by total small RNA isolation. RNA-seq analysis of small RNAs revealed a signature of differentially and exclusively packaged RNAs that distinguish NCD from B-ALL. The plasma EVs contained a heterogeneous mixture of miRNAs and fragments of lncRNA, tRNA and mRNA. Transcripts packaged in B-ALL EVs included those involved in the negative regulation of the cell cycle, suggesting that B-ALL cells are using EVs to discard genes that control growth. In contrast, NCD EVs carried genes representative of multiple organs, including brain, muscle, and epithelial cells. Overall, this signature could potentially be used to monitor B-ALL disease burden in pediatric B-ALL patients using blood draws instead of invasive bone marrow aspirates.

To understand the mechanism of RNA packaged from cells into EVs, we cultured pediatric B-ALL primary lymphoblasts and immortalized cell lines from which we isolated EVs. Based on transmission electron microscopy and nanoparticle tracking analyses, we found the EVs from primary lymphoblasts and immortalized cell lines to be similar in morphology and size range. Secondly, using small RNASeq, we found that a large portion of identified RNA cargo are currently unannotated or uncharacterized. Of all annotated RNA, miRNAs make up the major category, followed by mRNA, lncRNA and tRNA. We also found that pediatric B-ALL cells package RNAs into EVs using non-selective (loading) and non-selective (sorting) mechanisms.

The packaged miRNAs were both mature and precursor miRNAs, while the mRNAs, lncRNA and tRNAs were fragments and possibly full-length transcripts in a few cases. We also found motifs shared by a subset of RNAs exclusive to cells or EVs. Furthermore, we found that some of these transcripts are known to be linked to pediatric B-ALL disease onset and progression.

A comparison of RNA species packaged from pediatric B-ALL lymphoblasts into EVs, and differentially packaged RNA between NCD and pediatric B-ALL (**Fig 7.1**) showed that some DP RNAs in B-ALL plasma EVs were B cell-derived. Specifically, 49.6 % of the plasma EVs (481) are B cell derived. This is in line with previous knowledge as plasma contains a heterogenous mix of cells (different types of leukocytes including lymphocytes; erythrocytes and platelets), all of which secrete their EVs into blood. Based on this, it is a reasonable expectation that plasma contains EVs from numerous cell sources including B cells. This was found by researchers such as Muller and colleagues⁷¹³. Therefore, plasma differentially packaged EV signature of pediatric B-ALL consists of EVs from different blood cells; however, approximately half is from B-ALL cells. This may be because in B-ALL abnormally developed lymphoblasts crowd-out other cell types, making them the majority of cells, and thus their EVs are highly abundant in plasma. The 481 B cell specific RNAs differentially packaged in plasma EVs include lncRNAs such as TAGAP-AS1, LINC01446, LINC02755 and SH3TC2-DT and mRNAs such as FAM234B, NFYB, HLA-B, IGH, HDAC9 and ATAD5.

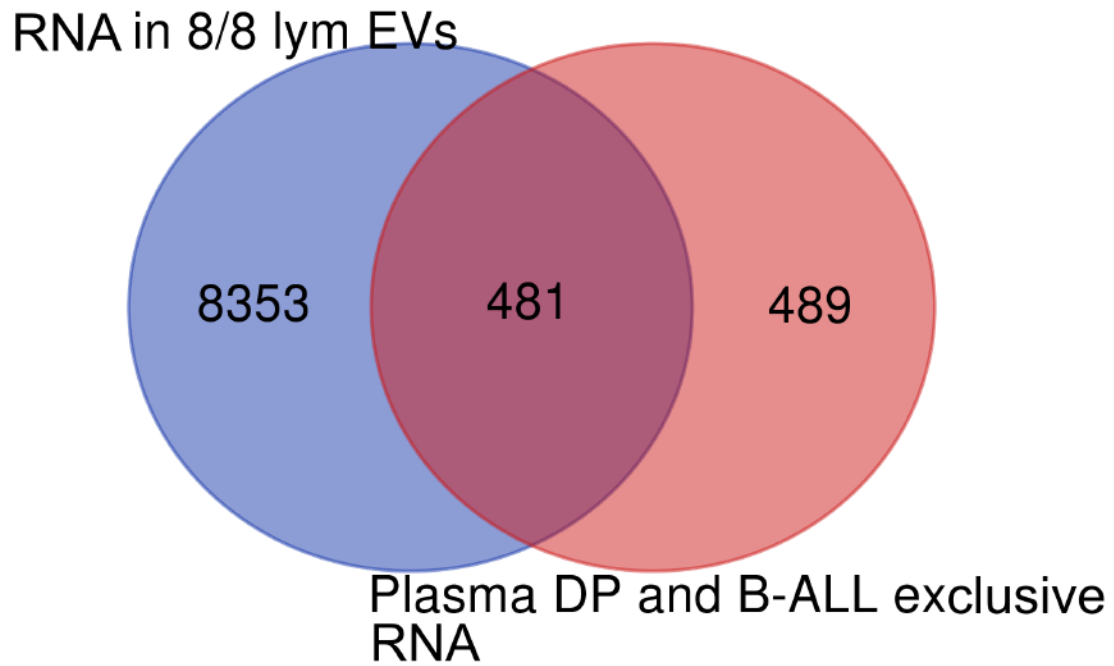


Fig. 7.1. B cell specific RNAs packaged into plasma EVs. Venn diagram comparing all RNAs present in 8/8 lymphoblast EVs and differentially packaged ($p < 0.05$) and B-ALL exclusive RNAs in plasma B-ALL EVs.

Others are miRNAs such as miR-3135B, miR-376C, and miR-193A; and unclassified RNA transcripts such as – LOC124904138, LOC105373153 and LOC105372047.

Finally, since we are interested in EVs as a source of biomarkers, we wanted to analyze key pre-analytical variables. The comparison between non-Streck treatment and Streck reagent showed that the EV size did not differ between collection conditions for blood plasma (EDTA anticoagulant) or urine samples, while the concentration did not differ between collection conditions for urine but differed for blood samples. Based on the total RNA and miRNA amounts obtained from the plasma (1ml) and urine (4ml) EVs, there was not enough material obtained from Streck collection tubes samples to adequately assess the level of all platelet and non-urine associated RNA originally included in our study. However, with the available RNA amounts, qRT-PCR analysis of selected miRNA transcripts showed that EDTA tubes contained more platelet activation associated miRNAs than Streck collected samples. There were reduced amounts of EV miRNA and total small RNA for urine in Streck-collected urine compared to EDTA tubes. Therefore, this raises questions about the effect of Streck on biofluid EVs. Overall, we can cautiously conclude that Streck preservative minimally affects EV concentration, and platelet RNA presence in EVs from blood plasma and urine. However, more work needs to be done to determine if increasing the volume of plasma or urine will be sufficient to allow the use of Streck tubes.

7.2 Future directions

7.2.1 Exploring EVs as biomarkers for monitoring pediatric B-ALL

Having identified a small RNA signature of pediatric B-ALL blood plasma, the next steps will include validating the top genes (chosen based on adjusted p-value <0.05 and fold change in each

category). The top genes can be experimentally validated using conventional qPCR; or digital droplet PCR (ddPCR) which can allow for absolute quantification, does not require standard curve, has a higher sensitivity and better resistance to PCR inhibitors. Using ddPCR would maximize available samples, which has been a limiting factor in our study using pediatric B-ALL patient plasma. This validation is important for correlating the RNASeq data to *in vivo* gene expression and packaging levels.

Next, a cross-sectional study could be done to explore the potential of pediatric B-ALL plasma EVs as a less invasive way to monitor MRD. Pediatric patient samples should be collected from peripheral blood, bone marrow and cerebrospinal fluid. EVs from these biofluids will then be quantified (mode and median diameter and concentration) followed by using small RNASeq and qPCR for EV small RNA cargo analysis. The cargo from control to disease will be compared to determine if the EV small RNA signature at sanctuary sites (BM and CNS) matches the signature in peripheral blood.

Additionally, a longitudinal study that involves collecting samples from pediatric B-ALL patients at different time points (diagnosis, remission, first and second relapse) would be valuable. This study would investigate the potential of EV size and concentration for monitoring disease burden and patient response to therapy; as well as if the EV RNA signature reflects disease burden in a way that allows it to be used as a biomarker. EV size and concentration can be analyzed longitudinally to ascertain if EV size and concentration changes longitudinally in pediatric B-ALL. Small RNASeq, targeted sequencing, and qPCR or ddPCR would also be used to analyze EV small RNA cargo profile longitudinally.

In addition, it is important to determine if these signatures are EV isolation technique dependent. For this purpose, other popular isolation techniques such as PEG-based ExoQuick, SEC-based

iZON columns, antibody- and microfluidics-based technique, could be compared, for EVs in PB. The EV RNA cargo profile should be compared using RNASeq and qPCR or ddPCR.

In addition to these objectives, during EV isolation from samples (using the mirVana optimized protocol from our lab), the organic layer of the RNA isolations can be kept and used for protein isolation. Alternatively, RNA, DNA and protein can simultaneously be isolated using kits such as the RNA/DNA/Protein purification plus kits (Norgen Biotek Corp, cat #47700) or the NucleoSpin TriPrep (Takara Bio, cat #:740966.50). The DNA can be used as an alternate input material for qPCR, while the proteins can be used for mass spectrometry analysis. The mass spectrometry analysis will be useful for getting the protein signature of pediatric B-ALL EVs, which could help with identifying proteins that could interact with the RNA cargo of EVs.

7.2.2 Elucidating the mechanism of RNA packaging into EVs.

The next step for the RNA packaging studies would be to validate the packaged RNA transcripts using qPCR or ddPCR. Then the miRNA packaging motifs should be validated using crosslinking, ligation and sequencing of hybrids (CLASH) to identify direct miRNA-target pairs⁷¹⁴. CLASH will allow for the identification of which mRNAs within the EVs interacts with the miRNAs packaged. Next, the RBPs associated with the packaged miRNAs can be identified using enhanced crosslinking and immunoprecipitation (eCLIP-seq)⁷¹⁵ for mapping the binding sites of RBPs on target RNA.

Next, mRNA, lncRNA and tRNA packaging mechanisms will need to be explored, starting by developing new tools to study the packaging mechanisms for the RNA biotypes. Specifically for mRNA, the first step would be determining if RBPs are important for mRNA packaging into EVs. This may be done by modifying previously established cell lines used in the course of this thesis,

based on an RNA-centric versus protein-centric approach⁷¹⁶. An RNA-centric approach starts with the RNA cargo of interest and uses that RNA to study the associated proteins, using methods such as RNA biotinylation for coupling to streptavidin which will allow for RNA pulldowns, and protein microarray, a high-throughput method for large scale tracking of protein interactions and functions. Conversely, protein-centric approach starts with a protein of interest and uses tools such as CLIP-seq and eCLIP-seq and characterize the proteins' interactions with RNA.

An additional approach to answering whether RNA transcripts are all fragments or full-length transcripts will be to use long-read sequencing using the PacBio system or mim-tRNASeq⁷¹⁷.

7.3 Significance

Using miRNAs as an analyte for less invasive monitoring of diseases including pediatric B-ALL, though promising, has not advanced much due to limitations in the current analytical tools available. However, when encapsulated within EVs, miRNAs, mRNAs and other RNA transcripts are more protected from degradation by nucleases than non-vesicle associated cell free RNA^{718,719}. EV RNA can also remain very stable for years when stored at temperatures including 4°C, making them easier to store and transport⁷²⁰⁻⁷²². Therefore, the EV small RNA signature of pediatric B-ALL plasma represents a potential new source of biomarkers for less invasive and real-time monitoring of disease burden and patient therapy response. This is significant for pediatric B-ALL patients who undergo multiple invasive procedures to monitor disease burden and track therapy response. This signature represents the future of liquid biopsy, which can improve patient outcomes, especially for patients who suffer relapse or take longer to reach remission.

For cancer, extracellular vesicle small RNA signature supplies new knowledge that could potentially lead to the update of the famous 'hallmarks of cancer' reviews^{723,724}, which explained

the molecular bases of cancers based on the molecular profile, and subsequent source of biomarkers. EVs may be the missing link that explains certain ‘not fully explained’ phenomena in cancer, such as drug resistance, paraneoplastic syndromes of the nervous system, unexplained relapse, spontaneous cancer regression and aspects of metastasis. Closer to home is the potential of using pediatric B-ALL EVs for less invasive monitoring of patients. Patients from rural NL would not need to come to St. John’s for BM biopsy and spinal taps. Instead, their blood plasma samples can be collected in their local health centre and shipped to St. John’s via cold chain storage. The samples can then be processed at the Janeway diagnostic laboratory, where these frozen samples can then be thawed, and EVs isolated using isolation techniques optimized during my Ph. D work. The RNA signature can then be analyzed using qPCR or targeted RNA-seq. Additionally, the NTA technique is simple enough that it can be integrated into clinical practice, with the EV quantification done on the thawed samples. Taken together, this study can offer new solutions to the healthcare problem for pediatric cancer patients in NL and potentially other parts of Canada.

The EV RNA cargo packaging hypothesis generated in this paper has implications for the EV field, biochemistry, and cancer. For the EV field, this cargo packaging hypothesis may fill the gap about the unknowns in the field. Knowing how RNA biotypes and which one gets packaged allows us to study disease pathology at a deeper level. Furthermore, understanding packaging mechanisms makes using EVs as conduits to deliver drugs and genes possible. Without a non-immunogenic delivery vector, the gene therapy field has not advanced as much as it could have. EVs represent a non-immunogenic particle that could potentially be loaded with nucleic material and targeted to specific sites. Therefore, knowing the packaging mechanism of RNA into EVs can be leveraged upon for engineering or loading of EVs with genetic material (such as RNA). EV loading with

RNA, coupled with knowledge of EV uptake and cargo delivery into recipients will facilitate use of EVs as non-viral vectors for gene therapy. Loading of EVs with nucleic acid, as gene therapy vectors is reviewed by different research groups^{725,726}. Another potential application could be the priming of environments, such as the BM, skin, heart and pancreas, with donor EVs, to prevent anti-graft inflammatory responses^{727,728}.

The study regarding EDTA vs Streck blood collection tubes raise questions about blood collection for clinical and research purposes, and the effect that the Streck preservative may have on biofluid samples collected for EV isolation purposes. With this knowledge, it may be time to rethink blood collection in clinics for diagnostics, therapeutics, prognostics, and biobanking purposes, if the pre-analytical processes including anticoagulation and platelet activation prevention pre-treatment can affect downstream uses including EV isolation. Overall, our study with EVs in B-ALL and EVs collected in various preprocessing containers contributes vital knowledge to EV research, biochemistry, and cancer biology.

Chapter 8: References

1. Kindt, T. J., Goldsby, R. A. & Osborne, B. A. Immunology. in *Kuby Immunology* (2019).
2. Alam, R. & Gorska, M. 3. Lymphocytes. in *Journal of Allergy and Clinical Immunology* vol. 111 (2003).
3. Gitlin, A. D. & Nussenzweig, M. C. Immunology: Fifty years of B lymphocytes. *Nature* **517**, 139–141 (2015).
4. Cooper, M. D., Peterson, R. D. A. & Good, R. A. Delineation of the thymic and bursal lymphoid systems in the chicken. *Nature* **205**, 143–146 (1965).
5. Eggenhofer, E., Luk, F., Dahlke, M. H. & Hoogduijn, M. J. The life and fate of mesenchymal stem cells. *Front. Immunol.* **5**, (2014).
6. Burnet, F. M. *The clonal selection theory of acquired immunity. The clonal selection theory of acquired immunity* (Vanderbilt University Press, 2011).
doi:10.5962/bhl.title.8281.
7. TALMAGE, D. W. Immunological specificity, unique combinations of selected natural globulins provide an alternative to the classical concept. *Science* **129**, 1643–8 (1959).
8. Owen, J. J. T., Cooper, M. D. & Raff, M. C. In vitro generation of B lymphocytes in mouse foetal liver, a mammalian ‘bursa equivalent’. *Nature* **249**, 361–363 (1974).
9. Cooper, M. D. A life of adventure in immunobiology. *Annu. Rev. Immunol.* **28**, 1–19 (2010).
10. Ruddle, N. H. Basics of Inducible Lymphoid Organs. *Curr. Top. Microbiol. Immunol.*

- 426, 1–19 (2020).
11. Bombardieri, M., Lewis, M. & Pitzalis, C. Ectopic lymphoid neogenesis in rheumatic autoimmune diseases. *Nat. Rev. Rheumatol.* **13**, 141–154 (2017).
 12. Gago da Graça, C., van Baarsen, L. G. M. & Mebius, R. E. Tertiary Lymphoid Structures: Diversity in Their Development, Composition, and Role. *J. Immunol.* **206**, 273–281 (2021).
 13. Hardy, R. & Hayakawa, K. B cell development pathways. *Annu. Rev. Immunol.* (2001) doi:10.1146/annurev.immunol.19.1.595.
 14. Osmond, D. G., Rolink, A. & Melchers, F. Murine B lymphopoiesis: Towards a unified model. *Immunol. Today* **19**, 65–68 (1998).
 15. Era, T. *et al.* Differentiation of growth signal requirement of B lymphocyte precursor is directed by expression of immunoglobulin. *EMBO J.* **10**, 337–342 (1991).
 16. Rodrigues, C. P., Shvedunova, M. & Akhtar, A. Epigenetic Regulators as the Gatekeepers of Hematopoiesis. *Trends Genet.* **37**, 125–142 (2021).
 17. Nutt, S. L. & Kee, B. L. The Transcriptional Regulation of B Cell Lineage Commitment. *Immunity* **26**, 715–725 (2007).
 18. Attaway, M., Chwat-Edelstein, T. & Vuong, B. Q. Regulatory Non-Coding RNAs Modulate Transcriptional Activation During B Cell Development. *Front. Genet.* **12**, (2021).
 19. Hagman, J. & Lukin, K. Transcription factors drive B cell development. *Curr. Opin. Immunol.* **18**, 127–134 (2006).

20. Pabst, R. Plasticity and heterogeneity of lymphoid organs. What are the criteria to call a lymphoid organ primary, secondary or tertiary? *Immunol. Lett.* **112**, 1–8 (2007).
21. Van De Pavert, S. A. & Mebius, R. E. New insights into the development of lymphoid tissues. *Nat. Rev. Immunol.* **10**, 664–674 (2010).
22. Randall, T. D. & Mebius, R. E. The development and function of mucosal lymphoid tissues: A balancing act with micro-organisms. *Mucosal Immunol.* **7**, 455–466 (2014).
23. Pilzecker, B. & Jacobs, H. Mutating for good: DNA damage responses during somatic hypermutation. *Front. Immunol.* **10**, (2019).
24. Yaari, G., Benichou, J. I. C., Vander Heiden, J. A., Kleinstein, S. H. & Louzoun, Y. The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).
25. Martin, A., Chahwan, R., Parsa, J. Y. & Scharff, M. D. Somatic Hypermutation: The Molecular Mechanisms Underlying the Production of Effective High-Affinity Antibodies. The Molecular Mechanisms Underlying the Production of Effective High-Affinity Antibodies. *Mol. Biol. B Cells Second Ed.* 363–388 (2014) doi:10.1016/B978-0-12-397933-9.00020-5.
26. Maul, R. W. & Gearhart, P. J. AID and somatic hypermutation. in *Advances in Immunology* (2010). doi:10.1016/S0065-2776(10)05006-6.
27. Birdsall, H. Adaptive Immunity: Antibodies and Immunodeficiencies. *Mand. Douglas, Bennett's Princ. Pract. Infect. Dis.* **1**, 39-49.e2 (2015).
28. Roche, P. A. & Furuta, K. The ins and outs of MHC class II-mediated antigen processing

- and presentation. *Nat. Rev. Immunol.* **15**, 203–216 (2015).
29. Marshall, J. S., Warrington, R., Watson, W. & Kim, H. L. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* **14**, (2018).
 30. Halperin, S. T., 'T Hart, B. A., Luchicchi, A. & Schenk, G. J. The Forgotten Brother: The Innate-like B1 Cell in Multiple Sclerosis. *Biomedicines* **10**, (2022).
 31. Montecino-Rodriguez, E. & Dorshkind, K. B-1 B Cell Development in the Fetus and Adult. *Immunity* **36**, 13–21 (2012).
 32. Haas, K. M., Poe, J. C., Steeber, D. A. & Tedder, T. F. B-1a and B-1b cells exhibit distinct developmental requirements and have unique functional roles in innate and adaptive immunity to *S. pneumoniae*. *Immunity* **23**, 7–18 (2005).
 33. Peng, B., Ming, Y. & Yang, C. Regulatory B cells: The cutting edge of immune tolerance in kidney transplantation review-Article. *Cell Death Dis.* **9**, (2018).
 34. Kaminski, D. A., Wei, C., Qian, Y., Rosenberg, A. F. & Sanz, I. Advances in human B cell phenotypic profiling. *Front. Immunol.* **3**, (2012).
 35. Wardemann, H., Boehm, T., Dear, N. & Carsetti, R. B-1a B cells that link the innate and adaptive immune responses are lacking in the absence of the spleen. *J. Exp. Med.* **195**, 771–780 (2002).
 36. Qorbani, A., Gao, G. & Dwyre, D. M. Polyclonal CD5⁺/CD19⁺ B1a lymphocytes after allogeneic stem cell transplantation: A potential diagnostic pitfall. *Autops. Case Reports* **10**, (2020).
 37. Zhang, X. Regulatory functions of innate-like B cells. *Cell. Mol. Immunol.* **10**, 113–121

- (2013).
38. Romero-Ramírez, S. *et al.* Innate-like B cell subsets during immune responses: Beyond antibody production. *J. Leukoc. Biol.* **105**, 843–856 (2019).
 39. Rothstein, T. L., Griffin, D. O., Holodick, N. E., Quach, T. D. & Kaku, H. Human B-1 cells take the stage. *Ann. N. Y. Acad. Sci.* **1285**, 97–114 (2013).
 40. Alugupalli, K. R. *et al.* B1b lymphocytes confer T cell-independent long-lasting immunity. *Immunity* **21**, 379–390 (2004).
 41. Mahajan, V. S. *et al.* B1a and B2 cells are characterized by distinct CpG modification states at DNMT3A-maintained enhancers. *Nat. Commun.* **12**, (2021).
 42. Chung, J. B., Sater, R. A., Fields, M. L., Erikson, J. & Monroe, J. G. CD23 defines two distinct subsets of immature B cells which differ in their responses to T cell help signals. *Int. Immunol.* **14**, 157–166 (2002).
 43. Martin, V. G. *et al.* Transitional B cells in early human B cell development - Time to revisit the paradigm? *Front. Immunol.* **7**, (2016).
 44. Suryani, S. & Tangye, S. G. Therapeutic implications of advances in our understanding of transitional B-cell development in humans. *Expert Rev. Clin. Immunol.* **6**, 765–775 (2010).
 45. Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science (80-.)*. **301**, 1374–1377 (2003).
 46. Allman, D. M., Ferguson, S. E., Lentz, V. M. & Cancro, M. P. Peripheral B cell maturation. II. Heat-stable antigen(hi) splenic B cells are an immature developmental

- intermediate in the production of long-lived marrow-derived B cells. *J. Immunol.* **151**, 4431–44 (1993).
47. Allman, D. M., Ferguson, S. E. & Cancro, M. P. Peripheral B cell maturation. I. Immature peripheral B cells in adults are heat-stable antigenhi and exhibit unique signaling characteristics. *J. Immunol.* **149**, 2533–40 (1992).
 48. Suryani, S. *et al.* Differential expression of CD21 identifies developmentally and functionally distinct subsets of human transitional B cells. *Blood* **115**, 519–529 (2010).
 49. Zhou, Y. *et al.* Transitional B cells involved in autoimmunity and their impact on neuroimmunological diseases. *J. Transl. Med.* **18**, (2020).
 50. Bemark, M. Translating transitions - How to decipher peripheral human B cell development. *J. Biomed. Res.* **29**, 264–284 (2015).
 51. Allman, D. *et al.* Resolution of Three Nonproliferative Immature Splenic B Cell Subsets Reveals Multiple Selection Points During Peripheral B Cell Maturation. *J. Immunol.* **167**, 6834–6840 (2001).
 52. Merrell, K. T. *et al.* Identification of Anergic B Cells within a Wild-Type Repertoire. *Immunity* **25**, 953–962 (2006).
 53. Cerutti, A., Cols, M. & Puga, I. Marginal zone B cells: Virtues of innate-like antibody-producing lymphocytes. *Nat. Rev. Immunol.* **13**, 118–132 (2013).
 54. Gary S. Firestein, Ralph C. Budd, ... James R. O'Dell. Kelley and Firestein's Textbook of Rheumatology. *Kelley Firestein's Textb. Rheumatol.* (2017) doi:10.1016/c2013-1-19259-3.

55. Scharenberg, A. M., Humphries, L. A. & Rawlings, D. J. Calcium signalling and cell-fate choice in B cells. *Nat. Rev. Immunol.* **7**, 778–789 (2007).
56. James, L. K. B cells defined by immunoglobulin isotypes. *Clin. Exp. Immunol.* (2022) doi:10.1093/cei/uxac091.
57. Akkaya, M., Kwak, K. & Pierce, S. K. B cell memory: building two walls of protection against pathogens. *Nat. Rev. Immunol.* **20**, 229–238 (2020).
58. Abdellah, Z. *et al.* Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
59. Ruiz-Orera, J., Messeguer, X., Subirana, J. A. & Alba, M. M. Long non-coding RNAs as a source of new peptides. *Elife* **3**, (2014).
60. S. Minchin & J. Lodge. Understanding biochemistry: Structure and function of nucleic acids. *Essays Biochem.* (2019).
61. de Yébenes, V. G., Bartolomé-Izquierdo, N. & Ramiro, A. R. Regulation of B-cell development and function by microRNAs. *Immunological Reviews* vol. 253 25–39 (2013).
62. Barnes, P. J. *et al.* Chapter 31 – Transcription Factors. in *Asthma and COPD* 373–380 (2009).
63. Rangarajan, A. A. Regulation of Gene Expression in Eukaryotes. *Genet. Fundam. Notes* 597–631 (2022) doi:10.1007/978-981-16-7041-1_13.
64. NESTLER, E. J. & HYMAN, S. E. Chapter 17: Regulation of gene expression. in *Neuropsychopharmacology: The Fifth Generation of Progress* vols 1–3 806–813 (2018).
65. Agrawal, S. & Ganley, A. R. D. The conservation landscape of the human ribosomal RNA

- gene repeats. *PLoS One* **13**, (2018).
66. Gonzalez, I. L. & Sylvester, J. E. Complete sequence of the 43-kb human ribosomal DNA repeat: Analysis of the intergenic spacer. *Genomics* **27**, 320–328 (1995).
 67. Henderson, A. S., Warburton, D. & Atwood, K. C. Location of ribosomal DNA in the human chromosome complement. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3394–3398 (1972).
 68. Tantravahi, R., Miller, D. A., Dev, V. G. & Miller, O. J. Detection of nucleolus organizer regions in chromosomes of human, chimpanzee, gorilla, orangutan and gibbon. *Chromosoma* **56**, 15–27 (1976).
 69. Schmickel, R. D. Quantitation of human ribosomal DNA: Hybridization of human DNA with ribosomal RNA for quantitation and fractionation. *Pediatr. Res.* **7**, 5–12 (1973).
 70. Stults, D. M., Killen, M. W., Pierce, H. H. & Pierce, A. J. Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome Res.* **18**, 13–18 (2008).
 71. Gibbons, J. G., Branco, A. T., Yu, S. & Lemos, B. Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. *Nat. Commun.* **5**, (2014).
 72. Parks, M. M. *et al.* Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci. Adv.* **4**, (2018).
 73. Falcon, K. T. *et al.* Dynamic regulation and requirement for ribosomal RNA transcription during mammalian development. *Proc. Natl. Acad. Sci. U. S. A.* **119**, (2022).
 74. Russell, J. & Zomerdijk, J. C. B. M. RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem. Sci.* **30**, 87–96 (2005).

75. Mandal, R. K. The Organization and Transcription of Eukaryotic Ribosomal RNA Genes. *Prog. Nucleic Acid Res. Mol. Biol.* **31**, 115–160 (1984).
76. Paule, M. R. SURVEY AND SUMMARY Transcription by RNA polymerases I and III. *Nucleic Acids Res.* **28**, 1283–1298 (2000).
77. Németh, A. *et al.* Initial genomics of the human nucleolus. *PLoS Genet.* **6**, (2010).
78. Yu, S. & Lemos, B. The long-range interaction map of ribosomal DNA arrays. *PLoS Genet.* **14**, (2018).
79. Torres-Machorro, A. L., Hernández, R., Cevallos, A. M. & López-Villaseñor, I. Ribosomal RNA genes in eukaryotic microorganisms: Witnesses of phylogeny? *FEMS Microbiol. Rev.* **34**, 59–86 (2010).
80. Kobayashi, T. & Ganley, A. R. D. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science (80-.).* **309**, 1581–1584 (2005).
81. Paredes, S. & Maggert, K. A. Ribosomal DNA contributes to global chromatin regulation. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 17829–17834 (2009).
82. Ganley, A. R. D. & Kobayashi, T. Ribosomal DNA and cellular senescence: New evidence supporting the connection between rDNA and aging. *FEMS Yeast Res.* **14**, 49–59 (2014).
83. Sinclair, D. A. & Guarente, L. Extrachromosomal rDNA circles - A cause of aging in yeast. *Cell* **91**, 1033–1042 (1997).
84. Derenzini, M. *et al.* Key role of the achievement of an appropriate ribosomal RNA complement for G1-S phase transition in H4-II-E-C3 rat hepatoma cells. *J. Cell. Physiol.*

- 202**, 483–491 (2005).
85. Deisenroth, C. & Zhang, Y. Ribosome biogenesis surveillance: Probing the ribosomal protein-Mdm2-p53 pathway. *Oncogene* **29**, 4253–4260 (2010).
 86. Boisvert, F. M., Van Koningsbruggen, S., Navascués, J. & Lamond, A. I. The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **8**, 574–585 (2007).
 87. Sirri, V., Urcuqui-Inchima, S., Roussel, P. & Hernandez-Verdun, D. Nucleolus: The fascinating nuclear body. *Histochem. Cell Biol.* **129**, 13–31 (2008).
 88. Ma, H. & Pederson, T. Nucleostemin: a multiplex regulator of cell-cycle progression. *Trends Cell Biol.* **18**, 575–579 (2008).
 89. Audas, T. E., Jacob, M. D. & Lee, S. Immobilization of Proteins in the Nucleolus by Ribosomal Intergenic Spacer Noncoding RNA. *Mol. Cell* **45**, 147–157 (2012).
 90. Zhang, L. F., Huynh, K. D. & Lee, J. T. Perinucleolar Targeting of the Inactive X during S Phase: Evidence for a Role in the Maintenance of Silencing. *Cell* **129**, 693–706 (2007).
 91. Gottlieb, S. & Esposito, R. E. A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. *Cell* **56**, 771–776 (1989).
 92. Krahn, N., Fischer, J. T. & Söll, D. Naturally Occurring tRNAs With Non-canonical Structures. *Front. Microbiol.* **11**, (2020).
 93. W.Pelley, J. 17 - Protein Synthesis and Degradation. in
doi:<https://www.sciencedirect.com/science/article/pii/B9780323074469000179>.
 94. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**, 234–237 (1969).

95. Weinmann, R. & Roeder, R. G. Role of DNA dependent RNA polymerase III in the transcription of the tRNA and 5S RNA genes. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 1790–1794 (1974).
96. HOSSENLOPP, P., CHAMBON, P. & WELLS, D. Animal DNA-Dependent RNA Polymerases: Partial Purification and Properties of Three Classes of RNA Polymerases from Uninfected and Adenovirus-Infected HeLa Cells. *Eur. J. Biochem.* **58**, 237–251 (1975).
97. Rubio, M. A. T. & Hopper, A. K. tRNA travels from the cytoplasm to organelles Mary. *Wiley Interdiscip. Rev. RNA* **2**, 802–817 (2011).
98. Rolle, K. *et al.* The sequence and structure determine the function of mature human miRNAs. *PLoS One* **11**, (2016).
99. Olena, A. F. & Patton, J. G. Genomic organization of microRNAs. *J. Cell. Physiol.* **222**, 540–545 (2010).
100. Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L. & Bradley, A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* **14**, 1902–1910 (2004).
101. Annese, T., Tamma, R., De Giorgis, M. & Ribatti, D. microRNAs Biogenesis, Functions and Role in Tumor Angiogenesis. *Front. Oncol.* **10**, (2020).
102. Ala, U. Competing Endogenous RNAs, Non-Coding RNAs and Diseases: An Intertwined Story. *Cells* **9**, (2020).
103. Friedman, R. C., Farh, K. K. H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs

- are conserved targets of microRNAs. *Genome Res.* (2009) doi:10.1101/gr.082701.108.
104. Bartel, D. P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* (2009) doi:10.1016/j.cell.2009.01.002.
 105. Iwasaki, Y. W., Siomi, M. C. & Siomi, H. PIWI-interacting RNA: Its biogenesis and functions. *Annu. Rev. Biochem.* **84**, 405–433 (2015).
 106. Ha, H. *et al.* A comprehensive analysis of piRNAs from adult human testis and their relationship with genes and mobile elements. *BMC Genomics* **15**, (2014).
 107. Wu, X. *et al.* The Biogenesis and Functions of piRNAs in Human Diseases. *Mol. Ther. - Nucleic Acids* **21**, 108–120 (2020).
 108. Carone, D. M. *et al.* A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. *Chromosoma* **118**, 113–125 (2009).
 109. Carone, D. M. *et al.* Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. *Chromosom. Res.* **21**, 49–62 (2013).
 110. Cao, F. *et al.* Dicer independent small RNAs associate with telomeric heterochromatin. *Rna* **15**, 1274–1281 (2009).
 111. Hari, R. & Parthasarathy, S. Prediction of coding and non-coding RNA. in *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics* vols 1–3 230–240 (2018).
 112. Mabin, J. W., Lewis, P. W., Brow, D. A. & Dvinge, H. Human spliceosomal snRNA sequence variants generate variant spliceosomes. *Rna* **27**, 1186–1203 (2021).
 113. Valadkhan, S. & Gunawardane, L. S. Role of small nuclear RNAs in eukaryotic gene

- expression. *Essays Biochem.* **54**, 79–90 (2013).
114. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **12**, 861–874 (2011).
115. Liang, J. *et al.* Small Nucleolar RNAs: Insight Into Their Function in Cancer. *Front. Oncol.* **9**, (2019).
116. Reichow, S. L., Hamma, T., Ferré-D'Amaré, A. R. & Varani, G. The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.* **35**, 1452–1464 (2007).
117. Huang, Z. hao, Du, Y. ping, Wen, J. tao, Lu, B. feng & Zhao, Y. snoRNAs: functions and mechanisms in biological processes, and roles in tumor pathophysiology. *Cell Death Discov.* **8**, (2022).
118. Fu, H. *et al.* Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett.* **583**, 437–442 (2009).
119. Thompson, D. M. & Parker, R. Stressing Out over tRNA Cleavage. *Cell* **138**, 215–219 (2009).
120. Thompson, D. M., Lu, C., Green, P. J. & Parker, R. tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *Rna* **14**, 2095–2103 (2008).
121. Tao, E. W. *et al.* A specific tRNA half, 5'tiRNA-His-GTG, responds to hypoxia via the HIF1 α /ANG axis and promotes colorectal cancer progression by regulating LATS2. *J. Exp. Clin. Cancer Res.* **40**, (2021).
122. Lee, Y. S., Shibata, Y., Malhotra, A. & Dutta, A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* **23**, 2639–2649 (2009).
123. Li, S., Xu, Z. & Sheng, J. tRNA-derived small RNA: A novel regulatory small non-coding

- RNA. *Genes (Basel)*. **9**, (2018).
124. Saikia, M. & Hatzoglou, M. The many virtues of tRNA-derived stress-induced RNAs (tiRNAs): Discovering novel mechanisms of stress response and effect on human health. *J. Biol. Chem.* **290**, 29761–29768 (2015).
 125. Cao, J., Cowan, D. B. & Wang, D. Z. tRNA-Derived Small RNAs and Their Potential Roles in Cardiac Hypertrophy. *Front. Pharmacol.* **11**, (2020).
 126. Lloret-Llinares, M., Mapendano, C. K., Martlev, L. H., Lykke-Andersen, S. & Jensen, T. H. Relationships between PROMPT and gene expression. *RNA Biol.* **13**, 6–14 (2016).
 127. Preker, P. *et al.* PROMoter uPstream Transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. *Nucleic Acids Res.* **39**, 7179–7193 (2011).
 128. Yu, D., Ma, X., Zuo, Z., Wang, H. & Meng, Y. Classification of transcription boundary-associated RNAs (TBARs) in animals and plants. *Front. Genet.* **9**, (2018).
 129. Al-Tobasei, R., Paneru, B. & Salem, M. Genome-wide discovery of long non-coding RNAs in rainbow trout. *PLoS One* **11**, (2016).
 130. Hu, S., Wu, J., Chen, L. & Shan, G. Signals from noncoding RNAs: Unconventional roles for conventional pol III transcripts. *Int. J. Biochem. Cell Biol.* **44**, 1847–1851 (2012).
 131. Statello, L., Guo, C. J., Chen, L. L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
 132. Sigova, A. A. *et al.* Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 2876–2881 (2013).

133. Zhang, X. *et al.* Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* **20**, (2019).
134. Dykes, I. M. & Emanuelli, C. Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA. *Genomics, Proteomics Bioinforma.* **15**, 177–186 (2017).
135. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).
136. Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **18**, 18–30 (2016).
137. Guo, C. J. *et al.* Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells. *Cell* **181**, 621-636.e22 (2020).
138. Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: The rosetta stone of a hidden RNA language? *Cell* **146**, 353–358 (2011).
139. Cesana, M. & Daley, G. Q. Deciphering the rules of ceRNA networks. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7112–7113 (2013).
140. Sen, R., Ghosal, S., Das, S., Balti, S. & Chakrabarti, J. Competing endogenous RNA: The key to posttranscriptional regulation. *Sci. World J.* **2014**, (2014).
141. Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
142. Cesana, M. *et al.* A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell* **147**, 947 (2011).
143. Ebert, M. S. & Sharp, P. A. MicroRNA sponges: Progress and possibilities. *Rna* **16**, 2043–

- 2050 (2010).
144. Kang, J. *et al.* Systematic Analysis of Competing Endogenous RNA Networks in Diffuse Large B-Cell Lymphoma and Hodgkin's Lymphoma. *Front. Genet.* **11**, (2020).
 145. Kartha, R. V. & Subramanian, S. Competing endogenous RNAs (ceRNAs): New entrants to the intricacies of gene regulation. *Front. Genet.* **5**, (2014).
 146. Ghafouri-Fard, S., Khoshbakht, T., Hussien, B. M., Taheri, M. & Jamali, E. The emerging role non-coding RNAs in B cell-related disorders. *Cancer Cell Int.* **22**, (2022).
 147. Fernandes, M., Marques, H., Teixeira, A. L. & Medeiros, R. Competitive endogenous RNA network involving miRNA and lncRNA in non-hodgkin lymphoma: Current advances and clinical perspectives. *Biomedicines* **9**, (2021).
 148. Xu, G. *et al.* The emerging roles of non-coding competing endogenous RNA in hepatocellular carcinoma. *Cancer Cell Int.* **20**, (2020).
 149. Kumar, V., Abbas, A. K., Aster, J. C. & Perkins, J. A. *Robbins and Cotran Pathologic Basis of Disease.* (Elsevier Inc., 2015).
 150. Hodgson, C. *Blood cancer in canada. Facts and Stats. 2016. Leukemia and Lymphoma Society of Canada.*
 151. Leukemia & Lymphoma Society. *Facts 2018-2019: UPDATED DATA ON BLOOD CANCERS.* (2019).
 152. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).

153. Grigoropoulos, N. F., Petter, R., Van 't Veer, M. B., Scott, M. a & Follows, G. a. Leukaemia update. Part 1: diagnosis and management. *BMJ* (2013) doi:10.1136/bmj.f1660.
154. Brown, P. A. *et al.* Acute lymphoblastic leukemia, version 2.2021. *JNCCN J. Natl. Compr. Cancer Netw.* **19**, 1079–1109 (2021).
155. Frishman-Levy, L. & Izraeli, S. Advances in understanding the pathogenesis of CNS acute lymphoblastic leukaemia and potential for therapy. *Br. J. Haematol.* **176**, 157–167 (2017).
156. Del Principe, M. I. *et al.* Central nervous system involvement in adult acute lymphoblastic leukemia: Diagnostic tools, prophylaxis, and therapy. *Mediterr. J. Hematol. Infect. Dis.* **6**, (2014).
157. Lenk, L., Alsadeq, A. & Schewe, D. M. Involvement of the central nervous system in acute lymphoblastic leukemia: opinions on molecular mechanisms and clinical implications based on recent data. *Cancer Metastasis Rev.* **39**, 173–187 (2020).
158. Dworzak, M. N. *et al.* AIEOP-BFM Consensus Guidelines 2016 for Flow Cytometric Immunophenotyping of Pediatric Acute Lymphoblastic Leukemia. *Cytom. Part B Clin. Cytom.* **94**, 82–93 (2018).
159. Pérez-Vera, P., Reyes-León, A. & Fuentes-Pananá, E. M. Signaling Proteins and Transcription Factors in Normal and Malignant Early B Cell Development. *Bone Marrow Res.* **2011**, 1–10 (2011).
160. Crist, W. *et al.* Poor prognosis of children with pre-B acute lymphoblastic leukemia is associated with the t(1;19)(q23;p13): a Pediatric Oncology Group study. *Blood* (1990)

doi:10.1182/blood.v76.1.117.bloodjournal761117.

161. Wafa, A., As'Sad, M., Liehr, T., Aljapawe, A. & Al Achkar, W. Childhood pre-B cell acute lymphoblastic leukemia with translocation t(1;19)(q21.1;p13.3) and two additional chromosomal aberrations involving chromosomes 1, 6, and 13: A case report. *J. Med. Case Rep.* (2017) doi:10.1186/s13256-017-1251-1.
162. Felice, M. S. *et al.* Prognostic impact of t(1;19)/ TCF3-PBX1 in childhood acute lymphoblastic leukemia in the context of Berlin-Frankfurt-Münster-based protocols. *Leuk. Lymphoma* **52**, 1215–1221 (2011).
163. Uckun, F. M. *et al.* Clinical significance of translocation t(1;19) in childhood acute lymphoblastic leukemia in the context of contemporary therapies: A report from the children's cancer group. *J. Clin. Oncol.* (1998) doi:10.1200/JCO.1998.16.2.527.
164. Ching-Hon Pui, W. E. E. Treatment of acute lymphoblastic leukemia. *N Engl J Med* . (2006).
165. Inaba, T. *et al.* Reversal of apoptosis by the leukaemia-associated E2A-HLF chimaeric transcription factor. *Nature* **382**, 541–544 (1996).
166. Hunger, S. P. Tyrosine kinase inhibitor use in pediatric Philadelphia chromosome-positive acute lymphoblastic anemia. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* (2011) doi:10.1182/asheducation-2011.1.361.
167. Raimondi, S. C. *et al.* Reassessment of the Prognostic Significance of Hypodiploidy in Pediatric Patients with Acute Lymphoblastic Leukemia. *Cancer* (2003)

doi:10.1002/cncr.11841.

168. Shurtleff, S. A. *et al.* TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia* (1995).
169. Heerema, N. A. *et al.* Prognostic impact of trisomies of chromosomes and 5 among children with acute lymphoblastic leukemia and high hyperdiploidy (> 50 chromosomes). *J. Clin. Oncol.* **18**, 1876–1887 (2000).
170. Harris, R. L., Harrison, C. J., Martineau, M., Taylor, K. E. & Moorman, A. V. Is trisomy 5 a distinct cytogenetic subgroup in acute lymphoblastic leukemia? *Cancer Genet. Cytogenet.* **148**, 159–162 (2004).
171. Look, A. T. Oncogenic transcription factors in the human acute leukemias. *Science* (80-.). **278**, 1059–1064 (1997).
172. Shiozawa, Y., Pedersen, E. A. & Taichman, R. S. GAS6/Mer axis regulates the homing and survival of the E2A/PBX1-positive B-cell precursor acute lymphoblastic leukemia in the bone marrow niche. *Exp. Hematol.* (2010) doi:10.1016/j.exphem.2009.11.002.
173. Mrózek, K., Harper, D. P. & Aplan, P. D. Cytogenetics and Molecular Genetics of Acute Lymphoblastic Leukemia. *Hematology/Oncology Clinics of North America* (2009) doi:10.1016/j.hoc.2009.07.001.
174. Van Dongen, J. J. M. *et al.* Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: Investigation of minimal residual disease in

- acute leukemia. *Leukemia* (1999) doi:10.1038/sj.leu.2401592.
175. Jain, S. & Abraham, A. BCR-ABL1-like B-acute lymphoblastic leukemia/lymphoma: A comprehensive review. *Arch. Pathol. Lab. Med.* **144**, 150–155 (2020).
176. Moorman, A. V. *et al.* Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. *Blood* (2003) doi:10.1182/blood-2003-04-1128.
177. Sutcliffe, M. J. *et al.* High concordance from independent studies by the Children's Cancer Group (CCG) and Pediatric Oncology Group (POG) associating favorable prognosis with combined trisomies 4, 10, and 17 in children with NCI Standard-Risk B-precursor Acute Lymphoblastic Leuk. *Leukemia* (2005) doi:10.1038/sj.leu.2403673.
178. Lilljebjörn, H. *et al.* Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat. Commun.* **7**, 11790 (2016).
179. Chiaretti, S., Zini, G. & Bassan, R. Diagnosis and Subclassification of Acute Lymphoblastic Leukemia. *Cit. Mediterr J Hematol Infect Dis* **6**, 2014073 (2014).
180. Arber, D. A. *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* (2016) doi:10.1182/blood-2016-03-643544.
181. Lilljebjörn, H. & Fioretos, T. New oncogenic subtypes in pediatric B-cell precursor acute lymphoblastic leukemia. *Blood* **130**, 1395–1401 (2017).
182. Brown, P. *et al.* Pediatric Acute Lymphoblastic Leukemia, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Cancer Netw.* **18**, 81–112

- (2020).
183. Pui, C. H. *et al.* Improved outcome for children with acute lymphoblastic leukemia: Results of Total Therapy Study XIII B at St Jude Children's Research Hospital. *Blood* (2004) doi:10.1182/blood-2004-04-1616.
 184. Duque-Afonso, J. *et al.* E2A-PBX1 remodels oncogenic signaling networks in B-cell precursor acute lymphoid leukemia. *Cancer Res.* **76**, 6937–6949 (2016).
 185. Hu, Y. *et al.* E2A-PBX1 exhibited a promising prognosis in pediatric acute lymphoblastic leukemia treated with the CCLG-ALL2008 protocol. *Onco. Targets. Ther.* **9**, 7219–7225 (2016).
 186. Juarez-Velazquez, M. R. *et al.* Genetic Markers in the Prognosis of Childhood Acute Lymphoblastic Leukemia. in *Clinical Epidemiology of Acute Lymphoblastic Leukemia - From the Molecules to the Clinic* (2013). doi:10.5772/54288.
 187. Rand, V. *et al.* Genomic characterization implicates iAMP21 as a likely primary genetic event in childhood B-cell precursor acute lymphoblastic leukemia. *Blood* **117**, 6848–6855 (2011).
 188. Bhojwani, D., Yang, J. J. & Pui, C. H. Biology of childhood acute lymphoblastic leukemia. *Pediatr. Clin. North Am.* **62**, 47–60 (2015).
 189. Moorman, A. V. *et al.* Risk-directed treatment intensification significantly reduces the risk of relapse among children and adolescents with acute lymphoblastic leukemia and intrachromosomal amplification of chromosome 21: A comparison of the MRC ALL97/99 and UKALL2003 trials. *J. Clin. Oncol.* (2013) doi:10.1200/JCO.2013.48.9377.

190. Paulsson, K. *et al.* The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Nat. Genet.* (2015) doi:10.1038/ng.3301.
191. Seeger, K. *et al.* TEL-AML1 fusion transcript in relapsed childhood acute lymphoblastic leukemia. The Berlin-Frankfurt-Münster Study Group. *Blood* **91**, 1716–22 (1998).
192. Liu, Y. F. *et al.* Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine* (2016) doi:10.1016/j.ebiom.2016.04.038.
193. Zhou, Y. *et al.* Advances in the molecular pathobiology of B-lymphoblastic leukemia. *Hum. Pathol.* (2012) doi:10.1016/j.humpath.2012.02.004.
194. Zuckerman, T. & Rowe, J. M. Pathogenesis and prognostication in acute lymphoblastic leukemia. *F1000Prime Rep.* (2014) doi:10.12703/P6-59.
195. Leukemia and Lymphoma Society. Facts and statistics: About blood cancers. *The Leukemia & Lymphoma Society of Canada* 1–19 <https://www.llscanada.org/disease-information/facts-and-statistics> (2018).
196. Luskin, M. R. & DeAngelo, D. J. Chimeric Antigen Receptor Therapy in Acute Lymphoblastic Leukemia Clinical Practice. *Curr. Hematol. Malig. Rep.* **12**, 370–379 (2017).
197. Dördelmann, M. *et al.* Prednisone response is the strongest predictor of treatment outcome in infant acute lymphoblastic leukemia. *Blood* **94**, 1209–1217 (1999).
198. Seibel, N. L. Treatment of acute lymphoblastic leukemia in children and adolescents: peaks and pitfalls. *Hematology Am. Soc. Hematol. Educ. Program* 374–380 (2008) doi:10.1182/asheducation-2008.1.374.

199. Wu, S. Y., Short, N. J., Nasr, L., Dabaja, B. S. & Fang, P. Q. Central Nervous System Prophylaxis and Treatment in Acute Leukemias. *Curr. Treat. Options Oncol.* (2022) doi:10.1007/s11864-022-01032-5.
200. Inaba, H., Greaves, M. & Mullighan, C. G. Acute lymphoblastic leukaemia. *The Lancet* (2013) doi:10.1016/S0140-6736(12)62187-4.
201. Truong, T. H. *et al.* Allogeneic Hematopoietic Stem Cell Transplantation for Children With Acute Lymphoblastic Leukemia: Shifting Indications in the Era of Immunotherapy. *Front. Pediatr.* **9**, (2021).
202. Quddus, F. F. *et al.* Glucocorticoid Receptors in Immunological Subtypes of Childhood Acute Lymphocytic Leukemia Cells: A Pediatric Oncology Group Study. *Cancer Res.* **45**, 6482–6486 (1985).
203. Tsai, S. Y. *et al.* Molecular interactions of steroid hormone receptor with its enhancer element: Evidence for receptor dimer formation. *Cell* **55**, 361–369 (1988).
204. Ray, A. & Prefontaine, K. E. Physical association and functional antagonism between the p65 subunit of transcription factor NF- κ B and the glucocorticoid receptor. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 752–756 (1994).
205. Tissing, W. J. E., Meijerink, J. P. P., den Boer, M. L. & Pieters, R. Molecular determinants of glucocorticoid sensitivity and resistance in acute lymphoblastic leukemia. *Leukemia* **17**, 17–25 (2003).
206. Zhuang, Y. *et al.* Reduced Dose Intensity of Daunorubicin During Remission Induction for Low-Risk Patients With Acute Lymphoblastic Leukemia: A Retrospective Cohort

- Study of the Chinese Children's Cancer Group. *Front. Oncol.* **12**, (2022).
207. Bassan, R. & Hoelzer, D. Modern therapy of acute lymphoblastic leukemia. *J. Clin. Oncol.* **29**, 532–543 (2011).
208. Hunger, S. P. *et al.* Improved Survival for Children and Adolescents With Acute Lymphoblastic Leukemia Between 1990 and 2005: A Report From the Children's Oncology Group. *J. Clin. Oncol.* **2012**, 207–208 (2012).
209. Stanulla, M. & Schrappe, M. Treatment of Childhood Acute Lymphoblastic Leukemia. *Semin. Hematol.* **46**, 52–63 (2009).
210. Pui, C. H., Robison, L. L. & Look, A. T. Acute lymphoblastic leukaemia. *Lancet* **371**, 1030–1043 (2008).
211. Trials, C. & Nccn, T. Leukemia Treatment Regimens: Acute Lymphoblastic Leukemia (ALL). *Cancer Therapy Advisor* 1–35
<https://www.cancertherapyadvisor.com/home/cancer-topics/hematologic-cancers/hematologic-cancers-treatment-regimens/leukemia-treatment-regimens-acute-lymphoblastic-leukemia-all/> (2020).
212. Terwilliger, T. & Abdul-Hay, M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer J.* **7**, 577 (2017).
213. O'Brien, S. *et al.* Outcome of adults with acute lymphocytic leukemia after second salvage therapy. *Cancer* **113**, 3186–3191 (2008).
214. Stam, R. W. *et al.* Targeting FLT3 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. *Blood* **106**, 2484–2490 (2005).

215. Lock, R. *et al.* Initial testing (stage 1) of the BH3 mimetic ABT-263 by the pediatric preclinical testing program. *Pediatr. Blood Cancer* **50**, 1181–1189 (2008).
216. Kinase, A. A. *et al.* Pediatric Preclinical Testing Program (PPTP) evaluation of the Aurora A Kinase Inhibitor MLN8237. 8237.
217. Aboukameel, A. *et al.* Superior Anti-Tumor Activity of the CD19-Directed Immunotoxin, SAR3419 to Rituximab in Non-Hodgkin’s Xenograft Animal Models: Preclinical Evaluation. *Blood* **110**, 2339–2339 (2007).
218. Zhukovsky, E. *et al.* an Fc-Engineered Anti-CD19 Monoclonal Antibody with In Vitro and In Vivo Efficacy against Lymphoma and Leukemia XmAb 5574 : cynomolgus monkey Legend : *Cell Transplant.* **6**, 5574 (2008).
219. Mengxuan, S., Fen, Z. & Runming, J. Novel Treatments for Pediatric Relapsed or Refractory Acute B-Cell Lineage Lymphoblastic Leukemia: Precision Medicine Era. *Front. Pediatr.* **10**, (2022).
220. Cooper, S. L. & Brown, P. A. Treatment of pediatric acute lymphoblastic leukemia. *Pediatric Clinics of North America* (2015) doi:10.1016/j.pcl.2014.09.006.
221. Locatelli, F., Schrappe, M., Bernardo, M. E. & Rutella, S. How i treat relapsed childhood acute lymphoblastic leukemia. *Blood* (2012) doi:10.1182/blood-2012-02-265884.
222. Lissat, A. *et al.* Other (Non-cns/testicular) extramedullary localizations of childhood relapsed acute lymphoblastic leukemia and lymphoblastic lymphoma—a report from the all-rez study group. *J. Clin. Med.* **10**, (2021).
223. Gaudichon, J. *et al.* Mechanisms of extramedullary relapse in acute lymphoblastic

- leukemia: Reconciling biological concepts and clinical issues. *Blood Rev.* **36**, 40–56 (2019).
224. Sikich, N., Soulodre, C., Sadasook, N., Sleeman, A. & Verhey, J. *Citation Health Quality Ontario. Minimal residual disease evaluation in childhood acute lymphoblastic leukemia: a clinical evidence review. Ont Health Technol Assess Ser. Ontario Health Technology Assessment Series* vol. 16
<https://www.hqontario.ca/Portals/0/Documents/evidence/reports/eba-mrd-1603-en.pdf>
(2016).
225. Lee, J. W. & Cho, B. Prognostic factors and treatment of pediatric acute lymphoblastic leukemia. *Korean J Pediatr* **60**, 129–137 (2017).
226. Cloos, J. *et al.* Comprehensive Protocol to Sample and Process Bone Marrow for Measuring Measurable Residual Disease and Leukemic Stem Cells in Acute Myeloid Leukemia. *J. Vis. Exp* 56386 (2018) doi:10.3791/56386.
227. Gajjar, A. *et al.* Traumatic lumbar puncture at diagnosis adversely affects outcome in childhood acute lymphoblastic leukemia. www.bloodjournal.org (2000).
228. Yao, Q., Bai, Y., Orfao, A. & Chim, C. S. Standardized Minimal Residual Disease Detection by Next-Generation Sequencing in Multiple Myeloma. *Front. Oncol.* **9**, 449 (2019).
229. Qin, X., Zhang, M.-Y. & Liu, W.-J. Application of minimal residual disease monitoring in pediatric patients with acute lymphoblastic leukemia. *Eur. Rev. Med. Pharmacol. Sci.* **22**, 6885–6895 (2018).

230. van Dongen, J. J. M. M., van der Velden, V. H. J. J., Brüggemann, M. & Orfao, A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: Need for sensitive, fast, and standardized technologies. *Blood* **125**, 3996–4009 (2015).
231. Spinella, J. F. *et al.* Mutational dynamics of early and late relapsed childhood ALL: rapid clonal expansion and long-term dormancy. *Blood Adv.* (2018)
doi:10.1182/bloodadvances.2017011510.
232. Bailey, L. C., Lange, B. J., Rheingold, S. R. & Bunin, N. J. Bone-marrow relapse in paediatric acute lymphoblastic leukaemia. *Lancet Oncol.* **9**, 873–883 (2008).
233. Shankland, K. R., Armitage, J. O. & Hancock, B. W. Non-Hodgkin lymphoma. *Lancet* (2012) doi:10.1016/S0140-6736(12)60605-9.
234. Nagpal, P. *et al.* Pediatric Hodgkin lymphoma- biomarkers, drugs, and clinical trials for translational science and medicine. *Oncotarget* (2016) doi:10.18632/oncotarget.11509.
235. Gallamini, A., Hutchings, M. & Ramadan, S. Clinical presentation and staging of Hodgkin lymphoma. *Seminars in Hematology* (2016) doi:10.1053/j.seminhematol.2016.05.005.
236. Iacobucci, I. & Mullighan, C. G. Genetic Basis of Acute Lymphoblastic Leukemia. *J. Clin. Oncol.* **35**, 975–983 (2017).
237. Hasserjian, R. P. Acute myeloid leukemia: advances in diagnosis and classification. *Int. J. Lab. Hematol.* **35**, 358–366 (2013).
238. Guièze, R. & Wu, C. J. Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood* **126**, 445–53 (2015).
239. Schmitz, R., Ceribelli, M., Pittaluga, S., Wright, G. & Staudt, L. M. Oncogenic

- mechanisms in Burkitt lymphoma. *Cold Spring Harb. Perspect. Med.* **4**, a014282 (2014).
240. Jabbour, E. & Kantarjian, H. Chronic myeloid leukemia: 2018 update on diagnosis, therapy and monitoring. *Am. J. Hematol.* **93**, 442–459 (2018).
241. Itzykson, R. *et al.* Prognostic Score Including Gene Mutations in Chronic Myelomonocytic Leukemia. *J. Clin. Oncol.* **31**, 2428–2436 (2013).
242. Cheson, B. D. *et al.* Recommendations for initial evaluation, staging, and response assessment of hodgkin and non-hodgkin lymphoma: The lugano classification. *Journal of Clinical Oncology* vol. 32 3059–3067 (American Society of Clinical Oncology, 2014).
243. Lobb, R. J. *et al.* Optimized exosome isolation protocol for cell culture supernatant and human plasma. doi:10.3402/jev.v4.27031.
244. LAULAGNIER, K. *et al.* Mast cell- and dendritic cell-derived exosomes display a specific lipid composition and an unusual membrane organization. *Biochem. J.* **380**, 161–171 (2004).
245. Tkach, M. & Théry, C. Communication by Extracellular Vesicles: Where We Are and Where We Need to Go. *Cell* **164**, 1226–1232 (2016).
246. Yáñez-Mó, M. *et al.* Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles* **4**, 1–60 (2015).
247. Whiteside, T. L. Profiling of plasma-derived extracellular vesicles cargo for diagnosis of pancreatic malignancy. *Ann. Transl. Med.* **5**, 501 (2017).
248. György, B. *et al.* Membrane vesicles, current state-of-the-art: Emerging role of extracellular vesicles. *Cell. Mol. Life Sci.* **68**, 2667–2688 (2011).

249. Baietti, M. F. *et al.* Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nat. Cell Biol.* **14**, 677–685 (2012).
250. Colombo, M. *et al.* Analysis of ESCRT functions in exosome biogenesis, composition and secretion highlights the heterogeneity of extracellular vesicles. *J. Cell Sci.* **126**, 5553–5565 (2013).
251. Cocucci, E. & Meldolesi, J. Ectosomes and exosomes: shedding the confusion between extracellular vesicles. *Trends Cell Biol.* **25**, 364–72 (2015).
252. Raposo, G. & Stoorvogel, W. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* (2013) doi:10.1083/jcb.201211138.
253. Wen, C. *et al.* Biological roles and potential applications of immune cell-derived extracellular vesicles. *J. Extracell. Vesicles* (2017) doi:10.1080/20013078.2017.1400370.
254. Sampaio, N. G., Cheng, L. & Eriksson, E. M. *The role of extracellular vesicles in malaria biology and pathogenesis. Malaria Journal* vol. 16 (2017).
255. Kowal, J. *et al.* Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E968–E977 (2016).
256. Kalra, H. *et al.* Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation. *PLoS Biol.* (2012) doi:10.1371/journal.pbio.1001450.
257. Pando, A., Reagan, J. L., Quesenberry, P. & Fast, L. D. Extracellular vesicles in leukemia. *Leuk. Res.* **64**, 52–60.
258. Cocucci, E., Racchetti, G. & Meldolesi, J. Shedding microvesicles: artefacts no more.

- Trends in Cell Biology* (2009) doi:10.1016/j.tcb.2008.11.003.
259. Abels, E. R. & Breakefield, X. O. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. doi:10.1007/s10571-016-0366-z.
260. Boyiadzis, M. & Whiteside, T. L. Information transfer by exosomes: A new frontier in hematologic malignancies. *Blood Rev.* **29**, 281–90 (2015).
261. van Dongen, H. M., Masoumi, N., Witwer, K. W. & Pegtel, D. M. Extracellular Vesicles Exploit Viral Entry Routes for Cargo Delivery. *Microbiol. Mol. Biol. Rev.* (2016) doi:10.1128/MMBR.00063-15.
262. Fitzgerald, W. *et al.* A System of Cytokines Encapsulated in ExtraCellular Vesicles. *Sci. Rep.* **8**, 1–11 (2018).
263. Van Deun, J. *et al.* The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. *J. Extracell. Vesicles* (2014) doi:10.3402/jev.v3.24858.
264. Greening, D. W., Xu, R., Ji, H., Tauro, B. J. & Simpson, R. J. A Protocol for Exosome Isolation and Characterization: Evaluation of Ultracentrifugation, Density-Gradient Separation, and Immunoaffinity Capture Methods. in 179–209 (Humana Press, New York, NY, 2015). doi:10.1007/978-1-4939-2550-6_15.
265. Théry, C. *et al.* Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. *Curr. Protoc. Cell Biol.* **3.22**, 1–29 (2006).
266. Brennan, K. *et al.* A comparison of methods for the isolation and separation of extracellular vesicles from protein and lipid particles in human serum. *Sci. Rep.* **10**, (2020).

267. Gupta, S. *et al.* An improvised one-step sucrose cushion ultracentrifugation method for exosome isolation from culture supernatants of mesenchymal stem cells. *Stem Cell Res. Ther.* **9**, (2018).
268. Talebjedi, B., Tasnim, N., Hoorfar, M., Mastromonaco, G. F. & De Almeida Monteiro Melo Ferraz, M. Exploiting Microfluidics for Extracellular Vesicle Isolation and Characterization: Potential Use for Standardized Embryo Quality Assessment. *Front. Vet. Sci.* **7**, (2021).
269. Furi, I., Momen-Heravi, F. & Szabo, G. Extracellular vesicle isolation: present and future. *Ann. Transl. Med.* (2017) doi:10.21037/atm.2017.03.95.
270. Ghosh, A. *et al.* Rapid isolation of extracellular vesicles from cell culture and biological fluids using a synthetic peptide with specific affinity for heat shock proteins. *PLoS One* (2014) doi:10.1371/journal.pone.0110443.
271. Lötvall, J. *et al.* Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. *J. Extracell. Vesicles* **3**, 26913 (2014).
272. Markowska, A., Pendergrast, R. S., Pendergrast, J. S. & Pendergrast, P. S. A novel method for the isolation of extracellular vesicles and RNA from urine. *J. Circ. Biomarkers* (2017) doi:10.1177/1849454417712666.
273. Szatanek, R. *et al.* The Methods of Choice for Extracellular Vesicles (EVs) Characterization. *Int. J. Mol. Sci.* (2017) doi:10.3390/ijms18061153.
274. Gardiner, C. *et al.* Techniques used for the isolation and characterization of extracellular

- vesicles: results of a worldwide survey'. *J. Extracell. Vesicles* (2016)
doi:10.3402/jev.v5.32945.
275. Chen, Y., Zhao, Y., Yin, Y., Jia, X. & Mao, L. Mechanism of cargo sorting into small extracellular vesicles. *Bioengineered* **12**, 8186–8201 (2021).
276. Mulcahy, L. A., Pink, R. C. & Carter, D. R. F. Routes and mechanisms of extracellular vesicle uptake. *Journal of Extracellular Vesicles* (2014) doi:10.3402/jev.v3.24641.
277. Van Niel, G., D'Angelo, G. & Raposo, G. Shedding light on the cell biology of extracellular vesicles. *Nature Reviews Molecular Cell Biology* (2018)
doi:10.1038/nrm.2017.125.
278. De Jong, O. G. *et al.* Cellular stress conditions are reflected in the protein and RNA content of endothelial cell-derived exosomes. *J. Extracell. Vesicles* (2012)
doi:10.3402/jev.v1i0.18396.
279. Li, Y. *et al.* Circular RNA is enriched and stable in exosomes: A promising biomarker for cancer diagnosis. *Cell Research* (2015) doi:10.1038/cr.2015.82.
280. Kucharzewska, P. & Belting, M. Emerging roles of extracellular vesicles in the adaptive response of tumour cells to microenvironmental stress. *Journal of Extracellular Vesicles* (2013) doi:10.3402/jev.v2i0.20304.
281. Ghosh, A. *et al.* Rapid isolation of extracellular vesicles from cell culture and biological fluids using a synthetic peptide with specific affinity for heat shock proteins. *PLoS One* (2014) doi:10.1371/journal.pone.0110443.
282. Mathivanan, S., Ji, H. & Simpson, R. J. Exosomes: Extracellular organelles important in

- intercellular communication. *Journal of Proteomics* (2010)
doi:10.1016/j.jprot.2010.06.006.
283. Kim, S. B. *et al.* Caspase-8 controls the secretion of inflammatory lysyl-tRNA synthetase in exosomes from cancer cells. *J. Cell Biol.* (2017) doi:10.1083/jcb.201605118.
284. Guix, F. X. *et al.* Tetraspanin 6: A pivotal protein of the multiple vesicular body determining exosome release and lysosomal degradation of amyloid precursor protein fragments. *Mol. Neurodegener.* (2017) doi:10.1186/s13024-017-0165-0.
285. Heijnen, B. H. F. G., Schiel, A. E., Fijnheer, R., Geuze, H. J. & Sixma, J. J. Activated Platelets Release Two Types of Membrane Vesicles : *Blood J.* (1999) doi:94(11):3791-9.
286. Morita, E. *et al.* Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO J.* (2007) doi:10.1038/sj.emboj.7601850.
287. Kalra, H. *et al.* Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation.
288. They, C. *et al.* Proteomic Analysis of Dendritic Cell-Derived Exosomes: A Secreted Subcellular Compartment Distinct from Apoptotic Vesicles. *J. Immunol.* (2001)
doi:10.4049/jimmunol.166.12.7309.
289. Li, J. *et al.* Microvesicles shed from microglia activated by the P2X7-p38 pathway are involved in neuropathic pain induced by spinal nerve ligation in rats. *Purinergic Signal.* (2017) doi:10.1007/s11302-016-9537-0.
290. Runz, S. *et al.* Malignant ascites-derived exosomes of ovarian carcinoma patients contain CD24 and EpCAM. *Gynecol. Oncol.* (2007) doi:10.1016/j.ygyno.2007.08.064.

291. Pfrieder, F. W. (Institut des Neurosciences Cellulaires et Integratives), Vitale, N. url.
292. Jenjaroenpun, P. *et al.* Characterization of RNA in exosomes secreted by human breast cancer cell lines using next-generation sequencing. *PeerJ* (2013) doi:10.7717/peerj.201.
293. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
294. Liang, L. G. *et al.* An integrated double-filtration microfluidic device for isolation, enrichment and quantification of urinary extracellular vesicles for detection of bladder cancer. *Sci. Rep.* (2017) doi:10.1038/srep46224.
295. Garcia-Martin, R. *et al.* MicroRNA sequence codes for small extracellular vesicle release and cellular retention. *Nature* (2021) doi:10.1038/s41586-021-04234-3.
296. Ghoshal, B., Bertrand, E. & Bhattacharyya, S. N. Non-canonical argonaute loading of extracellular vesicle-derived exogenous single-stranded miRNA in recipient cells. *J. Cell Sci.* **134**, (2021).
297. Shurtleff, M. J., Temoche-Diaz, M. M., Karfilis, K. V., Ri, S. & Schekman, R. Y-box protein 1 is required to sort microRNAs into exosomes in cells and in a cell-free reaction. *Elife* (2016) doi:10.7554/eLife.19276.
298. Temoche-Diaz, M. M. *et al.* Distinct mechanisms of microrna sorting into cancer cell-derived extracellular vesicle subtypes. *Elife* (2019) doi:10.7554/eLife.47544.
299. Nolte-'t Hoen, E. N. M. *et al.* Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res.* **40**, 9272–9285 (2012).

300. Thakur, B. K. *et al.* Double-stranded DNA in exosomes: A novel biomarker in cancer detection. *Cell Research* (2014) doi:10.1038/cr.2014.44.
301. Cai, J. *et al.* Transferred BCR/ABL DNA from K562 extracellular vesicles causes chronic myeloid leukemia in immunodeficient mice. *PLoS One* **9**, 1–11 (2014).
302. Kanemoto, S. *et al.* Multivesicular body formation enhancement and exosome release during endoplasmic reticulum stress. *Biochem. Biophys. Res. Commun.* **480**, 166–172 (2016).
303. O’Neill, C. P., Gilligan, K. E. & Dwyer, R. M. Role of extracellular vesicles (EVs) in cell stress response and resistance to cancer therapy. *Cancers (Basel)*. **11**, (2019).
304. Aman, Y. *et al.* Autophagy in healthy aging and disease. *Nat. Aging* **1**, 634–650 (2021).
305. Li, X., He, S. & Ma, B. Autophagy and autophagy-related proteins in cancer. *Mol. Cancer* **19**, (2020).
306. Xing, H., Tan, J., Miao, Y., Lv, Y. & Zhang, Q. Crosstalk between exosomes and autophagy: A review of molecular mechanisms and therapies. *J. Cell. Mol. Med.* **25**, 2297–2308 (2021).
307. Nishimura, T. & Tooze, S. A. Emerging roles of ATG proteins and membrane lipids in autophagosome formation. *Cell Discov.* **6**, (2020).
308. Liang, X. H. *et al.* Induction of autophagy and inhibition of tumorigenesis by beclin 1. *Nature* **402**, 672–676 (1999).
309. Ma, Y. *et al.* Exosomal transfer of cisplatin-induced mir-425-3p confers cisplatin resistance in NSCLC through activating autophagy. *Int. J. Nanomedicine* **14**, 8121–8132

- (2019).
310. Song, M. *et al.* Bystander autophagy mediated by radiation-induced exosomal MIR-7-5p in non-targeted human bronchial epithelial cells. *Sci. Rep.* **6**, (2016).
 311. Zhou, T. *et al.* α -synuclein accumulation in SH-SY5Y cell impairs autophagy in microglia by exosomes overloading miR-19a-3p. *Epigenomics* **11**, 1661–1677 (2019).
 312. Zhang, C., Gan, X., Liang, R. & Jian, J. Exosomes Derived From Epigallocatechin Gallate-Treated Cardiomyocytes Attenuated Acute Myocardial Infarction by Modulating MicroRNA-30a. *Front. Pharmacol.* **11**, (2020).
 313. Xu, Y. Q., Xu, Y. & Wang, S. H. Effect of exosome-carried MIR-30a on myocardial apoptosis in myocardial ischemia-reperfusion injury rats through regulating autophagy. *Eur. Rev. Med. Pharmacol. Sci.* **23**, 7066–7072 (2019).
 314. Colletti, M., Ceglie, D., Di Giannatale, A. & Nazio, F. Autophagy and Exosomes Relationship in Cancer: Friends or Foes? *Front. Cell Dev. Biol.* **8**, (2021).
 315. Nair, U. *et al.* SNARE proteins are required for macroautophagy. *Cell* **146**, 290–302 (2011).
 316. Hassanpour, M., Rezabakhsh, A., Rezaie, J., Nouri, M. & Rahbarghazi, R. Exosomal cargos modulate autophagy in recipient cells via different signaling pathways. *Cell Biosci.* **10**, (2020).
 317. Puri, C. *et al.* The RAB11A-Positive Compartment Is a Primary Platform for Autophagosome Assembly Mediated by WIPI2 Recognition of PI3P-RAB11A. *Dev. Cell* **45**, 114-131.e8 (2018).

318. Bhattacharya, S. *et al.* GAIP interacting protein C-Terminus regulates autophagy and exosome biogenesis of pancreatic cancer through metabolic pathways. *PLoS One* **9**, (2014).
319. Guo, H. *et al.* Atg5 Disassociates the V1V0-ATPase to Promote Exosome Production and Tumor Metastasis Independent of Canonical Macroautophagy. *Dev. Cell* **43**, 716-730.e7 (2017).
320. Murrow, L., Malhotra, R. & Debnath, J. ATG12-ATG3 interacts with Alix to promote basal autophagic flux and late endosome function. *Nat. Cell Biol.* **17**, 300–310 (2015).
321. Liu, J. *et al.* Distinct dasatinib-induced mechanisms of apoptotic response and exosome release in imatinib-resistant human chronic myeloid leukemia cells. *Int. J. Mol. Sci.* **17**, (2016).
322. Bhattacharya, S., McElhanon, K. E., Gushchina, L. V. & Weisleder, N. Role of phosphatidylinositol-4,5-bisphosphate 3-kinase signaling in vesicular trafficking. *Life Sci.* **167**, 39–45 (2016).
323. Kumar, D., Gupta, D., Shankar, S. & Srivastava, R. K. Biomolecular characterization of exosomes released from cancer stem cells: Possible implications for biomarker and treatment of cancer. *Oncotarget* **6**, 3280–3291 (2015).
324. Villarroya-Beltri, C. *et al.* ISGylation controls exosome secretion by promoting lysosomal degradation of MVB proteins. *Nat. Commun.* **7**, (2016).
325. Quesada-López, T. *et al.* The lipid sensor GPR120 promotes brown fat activation and FGF21 release from adipocytes. *Nat. Commun.* **7**, (2016).

326. Zou, W. *et al.* Exosome Release Is Regulated by mTORC1. *Adv. Sci.* **6**, (2019).
327. Perakis, S. & Speicher, M. R. Emerging concepts in liquid biopsies. *BMC Med.* **15**, (2017).
328. Lone, S. N. *et al.* Liquid biopsy: a step closer to transform diagnosis, prognosis and future of cancer treatments. *Mol. Cancer* **21**, (2022).
329. Siravegna, G., Marsoni, S., Siena, S. & Bardelli, A. Integrating liquid biopsies into the management of cancer. *Nat. Rev. Clin. Oncol.* **14**, 531–548 (2017).
330. Cheng, F., Su, L. & Qian, C. Circulating tumor DNA: A promising biomarker in the liquid biopsy of cancer. *Oncotarget* **7**, 48832–48841 (2016).
331. Pantel, K. & Alix-Panabières, C. Circulating tumour cells in cancer patients: Challenges and perspectives. *Trends Mol. Med.* **16**, 398–406 (2010).
332. Alix-Panabières, C. & Pantel, K. Liquid biopsy: from discovery to clinical implementation. *Mol. Oncol.* **15**, 1617–1621 (2021).
333. Taylor, C. *et al.* Peptide-affinity precipitation of extracellular vesicles and cell-free dna improves sequencing performance for the detection of pathogenic mutations in lung cancer patient plasma. *Int. J. Mol. Sci.* **21**, 1–19 (2020).
334. Roy, J. W. *et al.* A multiparametric extraction method for Vn96-isolated plasma extracellular vesicles and cell-free DNA that enables multi-omic profiling. *Sci. Rep.* **11**, 1–15 (2021).
335. Zhou, E. *et al.* Circulating extracellular vesicles are effective biomarkers for predicting response to cancer therapy. *EBioMedicine* **67**, (2021).

336. Wu, J. Y., Li, Y. J., Hu, X. Bin, Huang, S. & Xiang, D. X. Preservation of small extracellular vesicles for functional analysis and therapeutic applications: a comparative evaluation of storage conditions. *Drug Deliv.* **28**, 162–170 (2021).
337. Raposo, G., Stoorvogel, W., Graça Raposo & Willem Stoorvogel. Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.* **200**, 373–83 (2013).
338. Lässer, C. Identification and analysis of circulating exosomal microRNA in human body fluids. *Methods Mol. Biol.* **1024**, 109–128 (2013).
339. De Toro, J., Herschlik, L., Waldner, C. & Mongini, C. Emerging roles of exosomes in normal and pathological conditions: New insights for diagnosis and therapeutic applications. *Front. Immunol.* **6**, (2015).
340. Keller, S., Ridinger, J., Rupp, A. K., Janssen, J. W. G. & Altevogt, P. Body fluid derived exosomes as a novel template for clinical diagnostics. *J. Transl. Med.* **9**, (2011).
341. García-Silva, S., Gallardo, M. & Peinado, H. DNA-Loaded Extracellular Vesicles in Liquid Biopsy: Tiny Players With Big Potential? *Front. Cell Dev. Biol.* **8**, (2021).
342. LeBleu, V. S. & Kalluri, R. Exosomes as a Multicomponent Biomarker Platform in Cancer. *Trends in Cancer* **6**, 767–774 (2020).
343. Gilson, P., Merlin, J. L. & Harlé, A. Deciphering Tumour Heterogeneity: From Tissue to Liquid Biopsy. *Cancers (Basel)*. **14**, (2022).
344. König, L. *et al.* Elevated levels of extracellular vesicles are associated with therapy failure and disease progression in breast cancer patients undergoing neoadjuvant chemotherapy. *Oncoimmunology* **7**, (2018).

345. Van Dommelen, S. M. *et al.* Cetuximab treatment alters the content of extracellular vesicles released from tumor cells. *Nanomedicine* **11**, 881–890 (2016).
346. Vinik, Y. *et al.* Proteomic analysis of circulating extracellular vesicles identifies potential markers of breast cancer progression, recurrence, and response. *Sci. Adv.* **6**, (2020).
347. Aubertin, K. *et al.* Massive release of extracellular vesicles from cancer cells after photodynamic treatment or chemotherapy. *Sci. Rep.* **6**, (2016).
348. Torrano, V. *et al.* Vesicle-MaNiA: Extracellular vesicles in liquid biopsy and cancer. *Curr. Opin. Pharmacol.* **29**, 47–53 (2016).
349. Ismail, N. *et al.* Macrophage microvesicles induce macrophage differentiation and miR-223 transfer. *Blood* (2013) doi:10.1182/blood-2011-08-374793.
350. Sarkar, A., Mitra, S., Mehta, S., Raices, R. & Wewers, M. D. Monocyte derived microvesicles deliver a cell death message via encapsulated caspase-1. *PLoS One* (2009) doi:10.1371/journal.pone.0007140.
351. Ayre, D. C. *et al.* CD24 induces changes to the surface receptors of B cell microvesicles with variable effects on their RNA and protein cargo. *Sci. Rep.* (2017) doi:10.1038/s41598-017-08094-8.
352. Schiller, M. *et al.* During apoptosis HMGB1 is translocated into apoptotic cell-derived membrane vesicles. *Autoimmunity* (2013) doi:10.3109/08916934.2012.750302.
353. Coleman, L. G. *et al.* HMGB1/IL-1 β complexes in plasma microvesicles modulate immune responses to burn injury. *PLoS One* **13**, e0195335 (2018).
354. Bhatnagar, S. & Schorey, J. S. Exosomes released from infected macrophages contain

- Mycobacterium avium glycopeptidolipids and are proinflammatory. *J. Biol. Chem.* (2007) doi:10.1074/jbc.M702277200.
355. O'Neill, H. C. & Quah, B. J. C. Exosomes secreted by bacterially infected macrophages are proinflammatory. *Science Signaling* (2008) doi:10.1126/stke.16pe8.
356. Gasser, O. & Schifferli, J. A. Activated polymorphonuclear neutrophils disseminate anti-inflammatory microparticles by ectocytosis. *Blood* **104**, 2543–2548 (2004).
357. Raposo, G. *et al.* B lymphocytes secrete antigen-presenting vesicles. *J. Exp. Med.* **183**, 1161–72 (1996).
358. Teo, B. H. D. & Wong, S. H. MHC class II-associated invariant chain (Ii) modulates dendritic cells-derived microvesicles (DCMV)-mediated activation of microglia. *Biochem. Biophys. Res. Commun.* **400**, 673–678 (2010).
359. Szabó, G. T. *et al.* Critical role of extracellular vesicles in modulating the cellular effects of cytokines. *Cell. Mol. Life Sci.* (2014) doi:10.1007/s00018-014-1618-z.
360. Manček-Keber, M. *et al.* Extracellular vesicle-mediated transfer of constitutively active MyD88 L265P engages MyD88 wt and activates signaling. *Blood* **131**, 1720–1729 (2018).
361. Gutzeit, C. *et al.* Exosomes Derived from Burkitt's Lymphoma Cell Lines Induce Proliferation, Differentiation, and Class-Switch Recombination in B Cells. *J. Immunol.* **192**, 5852–62 (2014).
362. Haderk, F. *et al.* Tumor-derived exosomes modulate PD-L1 expression in monocytes. *Sci. Immunol.* (2017) doi:10.1126/sciimmunol.aah5509.
363. Lankar, D. *et al.* Dynamics of major histocompatibility complex class II compartments

- during B cell receptor-mediated cell activation. *J. Exp. Med.* **195**, 461–72 (2002).
364. Riolland, P., Lankar, D., Raposo, G., Bonnerot, C. & Hubert, P. BCR-bound antigen is targeted to exosomes in human follicular lymphoma B-cells. *Biol. Cell* (2008)
doi:10.1042/bc20060027.
365. Yeh, Y.-Y. *et al.* Characterization of CLL exosomes reveals a distinct microRNA signature and enhanced secretion by activation of BCR signaling. *Blood* **125**, 3297–305 (2015).
366. Ayre, D. C. *et al.* Dynamic regulation of CD24 expression and release of CD24-containing microvesicles in immature B cells in response to CD24 engagement. *Immunology* (2015) doi:10.1111/imm.12493.
367. Chan, C. J., Smyth, M. J. & Martinet, L. Molecular mechanisms of natural killer cell activation in response to cellular stress. *Cell Death and Differentiation* (2014)
doi:10.1038/cdd.2013.26.
368. Raulet, D. H. & Guerra, N. Oncogenic stress sensed by the immune system: role of natural killer cell receptors. *Nat. Rev. Immunol.* **9**, 568–580 (2009).
369. Pogge von Strandmann, E. *et al.* Human Leukocyte Antigen-B-Associated Transcript 3 Is Released from Tumor Cells and Engages the NKp30 Receptor on Natural Killer Cells. *Immunity* **27**, 965–974 (2007).
370. Reiners, K. S. *et al.* Soluble ligands for NK cell receptors promote evasion of chronic lymphocytic leukemia cells from NK cell anti-tumor activity. *Blood* (2013)
doi:10.1182/blood-2013-01-476606.

371. Hedlund, M., Nagaeva, O., Kargl, D., Baranov, V. & Mincheva-Nilsson, L. Thermal- and oxidative stress causes enhanced release of NKG2D ligand-bearing immunosuppressive exosomes in leukemia/lymphoma T and B cells. *PLoS One* **6**, (2011).
372. Hong, C.-S., Muller, L., Whiteside, T. L. & Boyiadzis, M. Plasma exosomes as markers of therapeutic response in patients with acute myeloid leukemia. *Front. Immunol.* **5**, 160 (2014).
373. Hong, C. S. *et al.* Circulating exosomes carrying an immunosuppressive cargo interfere with cellular immunotherapy in acute myeloid leukemia. *Sci. Rep.* (2017) doi:10.1038/s41598-017-14661-w.
374. Huan, J. *et al.* RNA trafficking by acute myelogenous leukemia exosomes. *Cancer Res.* (2013) doi:10.1158/0008-5472.CAN-12-2184.
375. Kumar, B. *et al.* Acute myeloid leukemia transforms the bone marrow niche into a leukemia-permissive microenvironment through exosome secretion. *Leukemia* (2018) doi:10.1038/leu.2017.259.
376. Hornick, N. I. *et al.* Serum Exosome MicroRNA as a minimally-invasive early biomarker of AML. *Sci. Rep.* (2015) doi:10.1038/srep11295.
377. Huan, J. *et al.* Coordinate regulation of residual bone marrow function by paracrine trafficking of AML exosomes. *Leukemia* **29**, 2285–95 (2015).
378. Zhao, C., Du, F., Zhao, Y., Wang, S. & Qi, L. Acute myeloid leukemia cells secrete microRNA-4532-containing exosomes to mediate normal hematopoiesis in hematopoietic stem cells by activating the LDOC1-dependent STAT3 signaling pathway. *Stem Cell Res.*

- Ther.* **10**, 384 (2019).
379. Corrado, C. *et al.* Exosome-mediated crosstalk between chronic myelogenous leukemia cells and human bone marrow stromal cells triggers an Interleukin 8-dependent survival of leukemia cells. *Cancer Lett.* (2014) doi:10.1016/j.canlet.2014.03.009.
380. Corrado, C. *et al.* Chronic myelogenous leukaemia exosomes modulate bone marrow microenvironment through activation of epidermal growth factor receptor. *J. Cell. Mol. Med.* (2016) doi:10.1111/jcmm.12873.
381. Milani, G. *et al.* Expression Profiling of Circulating Microvesicles Reveals Intercellular Transmission of Oncogenic Pathways. *Mol. Cancer Res.* (2017) doi:10.1158/1541-7786.MCR-16-0307.
382. Johnson, S. M. *et al.* Metabolic reprogramming of bone marrow stromal cells by leukemic extracellular vesicles in acute lymphoblastic leukemia. *Blood* (2016) doi:10.1182/blood-2015-12-688051.
383. Fei, F. *et al.* B-cell precursor acute lymphoblastic leukemia and stromal cells communicate through Galectin-3. *Oncotarget* (2015) doi:10.18632/oncotarget.3409.
384. Ghosh, A. K. *et al.* Circulating microvesicles in B-cell chronic lymphocytic leukemia can stimulate marrow stromal cells: Implications for disease progression. *Blood* (2010) doi:10.1182/blood-2009-09-242719.
385. Paggetti, J. *et al.* Exosomes released by chronic lymphocytic leukemia cells induce the transition of stromal cells into cancer-associated fibroblasts. *Blood* (2015) doi:10.1182/blood-2014-12-618025.

386. El-Saghir, J., Nassar, F., Tawil, N. & El-Sabban, M. ATL-derived exosomes modulate mesenchymal stem cells: Potential role in leukemia progression. *Retrovirology* (2016) doi:10.1186/s12977-016-0307-4.
387. Terpos, E., Christoulas, D. & Gavriatopoulou, M. Biology and treatment of myeloma related bone disease. *Metabolism*. **80**, 80–90 (2018).
388. Raimondi, L. *et al.* Involvement of multiple myeloma cell-derived exosomes in osteoclast differentiation. *Oncotarget* (2015) doi:10.18632/oncotarget.3830.
389. Raimondo, S. *et al.* Multiple myeloma-derived exosomes are enriched of amphiregulin (AREG) and activate the epidermal growth factor pathway in the bone microenvironment leading to osteoclastogenesis. *J. Hematol. Oncol.* (2019) doi:10.1186/s13045-018-0689-y.
390. Zhang, L. *et al.* Tumor-derived extracellular vesicles inhibit osteogenesis and exacerbate myeloma bone disease. *Theranostics* **9**, 196–209 (2019).
391. Moloudizargari, M. *et al.* The emerging role of exosomes in multiple myeloma. *Blood Rev.* **38**, 100595 (2019).
392. Wang, J. *et al.* Bone marrow stromal cell-derived exosomes as communicators in drug resistance in multiple myeloma cells. *Blood* (2014) doi:10.1182/blood-2014-03-562439.
393. Roccaro, A. M. *et al.* BM mesenchymal stromal cell-derived exosomes facilitate multiple myeloma progression. *J. Clin. Invest.* **123**, 1542–1555 (2013).
394. Yoshida, M. *et al.* miR-7977 inhibits the Hippo-YAP signaling pathway in bone marrow mesenchymal stromal cells. *PLoS One* **14**, e0213220 (2019).
395. Horiguchi, H. *et al.* Extracellular vesicle miR-7977 is involved in hematopoietic

- dysfunction of mesenchymal stromal cells via poly(rC) binding protein 1 reduction in myeloid neoplasms. *Haematologica* (2016) doi:10.3324/haematol.2015.134932.
396. Viola, S. *et al.* Alterations in acute myeloid leukaemia bone marrow stromal cell exosome content coincide with gains in tyrosine kinase inhibitor resistance. *British Journal of Haematology* (2016) doi:10.1111/bjh.13551.
397. Umezu, T., Ohyashiki, K., Kuroda, M. & Ohyashiki, J. H. Leukemia cell to endothelial cell communication via exosomal miRNAs. *Oncogene* (2013) doi:10.1038/onc.2012.295.
398. Tadokoro, H., Umezu, T., Ohyashiki, K., Hirano, T. & Ohyashiki, J. H. Exosomes derived from hypoxic leukemia cells enhance tube formation in endothelial cells. *J. Biol. Chem.* (2013) doi:10.1074/jbc.M113.480822.
399. Umezu, T. *et al.* Exosomal miR-135b shed from hypoxic multiple myeloma cells enhances angiogenesis by targeting factor-inhibiting HIF-1. *Blood* **124**, 3748–3757 (2014).
400. Mineo, M. *et al.* Exosomes released by K562 chronic myeloid leukemia cells promote angiogenesis in a src-dependent fashion. *Angiogenesis* (2012) doi:10.1007/s10456-011-9241-1.
401. Taverna, S. *et al.* Role of exosomes released by chronic myelogenous leukemia cells in angiogenesis. *Int. J. Cancer* (2012) doi:10.1002/ijc.26217.
402. Fang, Y. *et al.* PML–RARa modulates the vascular signature of extracellular vesicles released by acute promyelocytic leukemia cells. *Angiogenesis* (2016) doi:10.1007/s10456-015-9486-1.

403. Taverna, S. *et al.* Exosomal shuttling of miR-126 in endothelial cells modulates adhesive and migratory abilities of chronic myelogenous leukemia cells. *Mol. Cancer* (2014) doi:10.1186/1476-4598-13-169.
404. Namee, N. M. & O’Driscoll, L. Extracellular vesicles and anti-cancer drug resistance. *Biochim. Biophys. Acta - Rev. Cancer* **1870**, 123–136 (2018).
405. Head, D. *et al.* Effect of aggressive daunomycin therapy on survival in acute promyelocytic leukemia. *Blood* **86**, 1717–28 (1995).
406. Filipits, M. *et al.* Multidrug resistance-associated protein in acute myeloid leukemia: No impact on treatment outcome. *Clin. Cancer Res.* **3**, 1419–25 (1997).
407. Bouvy, C. *et al.* Transfer of multidrug resistance among acute myeloid leukemia cells via extracellular vesicles and their microRNA cargo. *Leuk Res* **62**, 70–76 (2017).
408. Bebawy, M. *et al.* Membrane microparticles mediate transfer of P-glycoprotein to drug sensitive cancer cells. *Leukemia* **23**, 1643–1649 (2009).
409. Aung, T. *et al.* Exosomal evasion of humoral immunotherapy in aggressive B-cell lymphoma modulated by ATP-binding cassette transporter A3. *Proc. Natl. Acad. Sci.* **108**, 15336–15341 (2011).
410. Oksvold, M. P. *et al.* Expression of B-Cell Surface Antigens in Subpopulations of Exosomes Released From B-Cell Lymphoma Cells. *Clin. Ther.* **36**, 847-862.e1 (2014).
411. Chapuy, B. *et al.* Intracellular ABC transporter A3 confers multidrug resistance in leukemia cells by lysosomal drug sequestration. *Leukemia* **22**, 1576–1586 (2008).
412. García-Romero, N. *et al.* DNA sequences within glioma-derived extracellular vesicles can

- cross the intact blood-brain barrier and be detected in peripheral blood of patients. *Oncotarget* **8**, 1416–1428 (2017).
413. Balusu, S. *et al.* Identification of a novel mechanism of blood–brain communication during peripheral inflammation via choroid plexus-derived extracellular vesicles. *EMBO Mol. Med.* (2016) doi:10.15252/emmm.201606271.
414. Caivano, A. *et al.* High serum levels of extracellular vesicles expressing malignancy-related markers are released in patients with various types of hematological neoplastic disorders. *Tumor Biol.* **36**, 9739–52 (2015).
415. Szczepanski, M. J., Szajnik, M., Welsh, A., Whiteside, T. L. & Boyiadzis, M. Blast-derived microvesicles in sera from patients with acute myeloid leukemia suppress natural killer cell function via membrane-associated transforming growth factor- β 1. *Haematologica* **96**, 1302–1309 (2011).
416. van Eijndhoven, M. A. J. *et al.* Plasma vesicle miRNAs for therapy response monitoring in Hodgkin lymphoma patients. *JCI Insight* (2016) doi:10.1172/jci.insight.89631.
417. Théry, C. *et al.* Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* **7**, 1535750 (2018).
418. Rutherford, S. C. *et al.* Extracellular vesicles in DLBCL provide abundant clues to aberrant transcriptional programming and genomic alterations. *Blood* (2018) doi:10.1182/blood-2017-12-821843.
419. Caivano, A. *et al.* MicroRNA-155 in serum-derived extracellular vesicles as a potential

- biomarker for hematologic malignancies - a short report. *Cell. Oncol.* (2017)
doi:10.1007/s13402-016-0300-x.
420. Manier, S. *et al.* Prognostic role of circulating exosomal miRNAs in multiple myeloma. *Blood* (2017) doi:10.1182/blood-2016-09-742296.
421. Reiners, K. S. *et al.* Extracellular vesicles released from chronic lymphocytic leukemia cells exhibit a disease relevant mRNA signature and transfer mRNA to bystander cells. *Haematologica* (2017) doi:10.3324/haematol.2016.153197.
422. Hansen, H. P. *et al.* Protrusion-guided extracellular vesicles mediate CD30 trans-signalling in the microenvironment of Hodgkin's lymphoma. *J. Pathol.* (2014)
doi:10.1002/path.4306.
423. De Luca, L. *et al.* Characterization and prognostic relevance of circulating microvesicles in chronic lymphocytic leukemia. *Leuk. Lymphoma* (2017)
doi:10.1080/10428194.2016.1243790.
424. Boysen, J. *et al.* Dynamics of microvesicle generation in B-cell chronic lymphocytic leukemia: Implication in disease progression. *Leukemia* (2017) doi:10.1038/leu.2016.217.
425. Krishnan, S. R. *et al.* Isolation of Human CD138+ Microparticles from the Plasma of Patients with Multiple Myeloma. *Neoplasia* (2016) doi:10.1016/j.neo.2015.11.011.
426. Yi, J. H., Kim, S. J. & Kim, W. S. Brentuximab vedotin: clinical updates and practical guidance. *Blood Res.* **52**, 243 (2017).
427. Gheldof, D. *et al.* Procoagulant activity of extracellular vesicles as a potential biomarker for risk of thrombosis and DIC in patients with acute leukaemia. *J. Thromb. Thrombolysis*

- (2017) doi:10.1007/s11239-016-1471-z.
428. Gheldof, D. *et al.* Thrombin generation assay and transmission electron microscopy: A useful combination to study tissue factor-bearing microvesicles. *J. Extracell. Vesicles* **2**, 19728 (2013).
 429. Gustafson, H. H., Holt-Casper, D., Grainger, D. W. & Ghandehari, H. Nanoparticle uptake: The phagocyte problem. *Nano Today* vol. 10 487–510 (2015).
 430. Hoshino, A. *et al.* Tumour exosome integrins determine organotropic metastasis. *Nature* (2015) doi:10.1038/nature15756.
 431. Caivano, A. *et al.* Extracellular vesicles in hematological malignancies: From biology to therapy. *International Journal of Molecular Sciences* (2017) doi:10.3390/ijms18061183.
 432. Burnouf, T., Agrahari, V. & Agrahari, V. Extracellular Vesicles As Nanomedicine: Hopes And Hurdles In Clinical Translation. *Int. J. Nanomedicine* **Volume 14**, 8847–8859 (2019).
 433. Tian, Y. *et al.* A doxorubicin delivery platform using engineered natural membrane vesicle exosomes for targeted tumor therapy. *Biomaterials* **35**, 2383–2390 (2014).
 434. Liang, G. *et al.* Engineered exosomes for targeted co-delivery of miR-21 inhibitor and chemotherapeutics to reverse drug resistance in colon cancer. *J. Nanobiotechnology* **18**, 10 (2020).
 435. Bellavia, D. *et al.* Interleukin 3- receptor targeted exosomes inhibit in vitro and in vivo Chronic Myelogenous Leukemia cell growth. *Theranostics* **7**, 1333–1345 (2017).
 436. Dong, Y. *et al.* Targeted blocking of miR328 lysosomal degradation with alkalized exosomes sensitizes the chronic leukemia cells to imatinib. *Appl. Microbiol. Biotechnol.*

- 103**, 9569–9582 (2019).
437. De Miguel, D. *et al.* Liposomes decorated with Apo2L/TRAIL overcome chemoresistance of human hematologic tumor cells. *Mol. Pharm.* **10**, 893–904 (2013).
438. De Miguel, D., Gallego-Lleyda, A., Anel, A. & Martinez-Lostao, L. Liposome-bound TRAIL induces superior DR5 clustering and enhanced DISC recruitment in histiocytic lymphoma U937 cells. *Leuk. Res.* (2015) doi:10.1016/j.leukres.2015.03.019.
439. Gehrman, U., Näslund, T. I., Hiltbrunner, S., Larssen, P. & Gabrielsson, S. Harnessing the exosome-induced immune response for cancer immunotherapy. *Seminars in Cancer Biology* (2014) doi:10.1016/j.semcancer.2014.05.003.
440. Hao, S., Bai, O., Yuan, J., Qureshi, M. & Xiang, J. Dendritic cell-derived exosomes stimulate stronger CD8⁺ CTL responses and antitumor immunity than tumor cell-derived exosomes. *Cell. Mol. Immunol.* **3**, 205–11 (2006).
441. Colino, J. & Snapper, C. M. Exosomes from Bone Marrow Dendritic Cells Pulsed with Diphtheria Toxoid Preferentially Induce Type 1 Antigen-Specific IgG Responses in Naive Recipients in the Absence of Free Antigen. *J. Immunol.* **177**, 3757–3762 (2006).
442. Théry, C. *et al.* Indirect activation of naïve CD4⁺ T cells by dendritic cell-derived exosomes. *Nat. Immunol.* **3**, 1156–1162 (2002).
443. Qazi, K. R., Gehrman, U., Jordö, E. D., Karlsson, M. C. I. & Gabrielsson, S. Antigen-loaded exosomes alone induce Th1-type memory through a B cell dependent mechanism. *Blood* (2009) doi:10.1182/blood-2008-04-153536.
444. Shen, C., Hao, S. G., Zhao, C. X., Zhu, J. & Wang, C. Antileukaemia immunity: Effect of

- exosomes against NB4 acute promyelocytic leukaemia cells. *J. Int. Med. Res.* (2011)
doi:10.1177/147323001103900305.
445. Yao, Y. *et al.* Dendritic cells pulsed with leukemia cell-derived exosomes more efficiently induce antileukemic immunities. *PLoS One* (2014) doi:10.1371/journal.pone.0091463.
446. Escudier, B. *et al.* Vaccination of metastatic melanoma patients with autologous dendritic cell (DC) derived-exosomes: Results of the first phase 1 clinical trial. *J. Transl. Med.* **3**, 10 (2005).
447. Bruserud, Ø., Glenjen, N., Rynningen, A. & Ulvestad, E. In vitro culture of human acute lymphoblastic leukemia (ALL) cells in serum-free media; a comparison of native ALL blasts, ALL cell lines and virus-transformed B cell lines. *Leuk. Res.* **27**, 455–464 (2003).
448. Joy, A. P. *et al.* Proteome profiling of extracellular vesicles captured with the affinity peptide Vn96: comparison of Laemmli and TRIzol© protein-extraction methods. *J. Extracell. Vesicles* (2018) doi:10.1080/20013078.2018.1438727.
449. Page, M. J. *et al.* The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **382**, (2021).
450. Kratz, A., Tomita, M. & Krishnan, A. GeNESiS: Gene network evolution simulation software. *BMC Bioinformatics* **9**, (2008).
451. Dweep, H., Gretz, N. & Sticht, C. MiRWalk database for miRNA-target interactions. *Methods Mol. Biol.* **1182**, 289–305 (2014).
452. Chen, Y. & Wang, X. MiRDB: An online database for prediction of functional microRNA targets. *Nucleic Acids Res.* **48**, D127–D131 (2020).

453. Bindea, G. *et al.* ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).
454. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
455. Andrews, S. & others. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
<http://www.bioinformatics.babraham.ac.uk/projects/> (2010).
456. Ewels Philip, Magnusson Måns, Lundin Sverker & Käller Max. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
457. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R22 (2009).
458. Heng, L. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
459. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
460. Rueda, A. *et al.* SRNAToolbox: An integrated collection of small RNA research tools. *Nucleic Acids Res.* **43**, W467–W473 (2015).
461. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
462. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) doi:10.1186/s13059-014-0550-8.

463. Sturn, A., Quackenbush, J. & Trajanoski, Z. Genesis: Cluster analysis of microarray data. *Bioinformatics* **18**, 207–208 (2002).
464. Shannon, P. *et al.* Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
465. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **16**, 284–287 (2012).
466. Luo, Weijun, B. & Cory. “Pathview: an R/Bioconductor package for pathway-based data integration and visualization.” *Bioinformatics* **29**, 1830–1831 (2013).
467. Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995 (2022).
468. Wickham, H. ggplot2: Elegant Graphics for Data Analysis - Bookreview. *J. Stat. Softw.* **35**, 1–3 (2010).
469. Liu, W. & Wang, X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol.* (2019)
doi:10.1186/s13059-019-1629-z.
470. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, (2015).
471. Bailey, T. L., Johnson, J., Grant, C. E. & Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **43**, W39–W49 (2015).
472. Namayandeh, S. M., Khazaei, Z., Najafi, M. L., Goodarzi, E. & Moslem, A. GLOBAL Leukemia in children 0-14 statistics 2018, incidence and mortality and human development index (HDI): GLOBOCAN sources and methods. *Asian Pacific J. Cancer*

- Prev.* **21**, 1487–1494 (2020).
473. Gaynon, P. S. *et al.* Long-term results of the children’s cancer group studies for childhood acute lymphoblastic leukemia 1983–2002: A Children’s Oncology Group Report. *Leukemia* **24**, 285–297 (2010).
474. McAnena, P. *et al.* Circulating microRNAs miR-331 and miR-195 differentiate local luminal a from metastatic breast cancer. *BMC Cancer* **19**, (2019).
475. Shukla, V. *et al.* Enumeration of deregulated miRNAs in liquid and tissue biopsies of cervical cancer. *Gynecol. Oncol.* **155**, 135–143 (2019).
476. Humplikova, L. *et al.* Expression of miR-15a and miR-16-1 in patients with chronic lymphocytic leukemia. *Biomed. Pap.* **157**, 284–293 (2013).
477. Srinivasan, S. *et al.* Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation. *Cell* **177**, 446–462.e16 (2019).
478. Brown, R. A. M. *et al.* Total RNA extraction from tissues for microRNA and target gene expression analysis: Not all kits are created equal. *BMC Biotechnol.* **18**, (2018).
479. Leshkowitz, D., Horn-Saban, S., Parmet, Y. & Feldmesser, E. Differences in microRNA detection levels are technology and sequence dependent. *Rna* **19**, 527–538 (2013).
480. Gao, C. *et al.* Clinical-biological characteristics and treatment outcomes of pediatric pro-B ALL patients enrolled in BCH-2003 and CCLG-2008 protocol: a study of 121 Chinese children. *Cancer Cell Int.* **19**, (2019).
481. Øbro, N. F. *et al.* Immunophenotype-defined sub-populations are common at diagnosis in childhood B-cell precursor acute lymphoblastic leukemia. *Leukemia* **25**, 1652–1657

- (2011).
482. Organista-Nava, J., Gómez-Gómez, Y., Illades-Aguilar, B. & Leyva-Vázquez, M. A. Regulation of the miRNA expression by TEL/AML1, BCR/ABL, MLL/AF4 and TCF3/PBX1 oncoproteins in acute lymphoblastic leukemia (Review). *Oncology Reports* vol. 36 1226–1232 (2016).
 483. Huen, K., Lizarraga, D., Kogut, K., Eskenazi, B. & Holland, N. Age-related differences in miRNA expression in Mexican-American newborns and children. *Int. J. Environ. Res. Public Health* **16**, (2019).
 484. Kleiveland, C. & Kleiveland, C. Peripheral blood mononuclear cells. in *The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models* 161–167 (2015).
doi:10.1007/978-3-319-16104-4_15.
 485. Cobaleda, C. & Sánchez-García, I. B-cell acute lymphoblastic leukaemia: Towards understanding its cellular origin. *BioEssays* vol. 31 600–609 (2009).
 486. Yang, Y. *et al.* MiR-128-2 inhibits common lymphoid progenitors from developing into progenitor B cells. *Oncotarget* **7**, 17520–17531 (2016).
 487. Kramer, N. J. *et al.* Altered lymphopoiesis and immunodeficiency in miR-142 null mice. *Blood* **125**, 3720–3730 (2015).
 488. Sun, X., Sit, A. & Feinberg, M. W. Role of miR-181 family in regulating vascular inflammation and immunity. *Trends in Cardiovascular Medicine* vol. 24 105–112 (2014).
 489. Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs Modulate Hematopoietic Lineage Differentiation. *Science (80-.)*. **303**, 83–86 (2004).

490. Mathelier, A. & Carbone, A. Large scale chromosomal mapping of human microRNA structural clusters. *Nucleic Acids Res.* **41**, 4392–4408 (2013).
491. Kordes, U., Krappmann, D., Heissmeyer, V., Ludwig, W. D. & Scheidereit, C. Transcription factor NF- κ B is constitutively activated in acute lymphoblastic leukemia cells. *Leukemia* **14**, 399–402 (2000).
492. Viatour, P., Merville, M. P., Bours, V. & Chariot, A. Phosphorylation of NF- κ B and I κ B proteins: Implications in cancer and inflammation. *Trends Biochem. Sci.* **30**, 43–52 (2005).
493. Zhang, H. *et al.* NF-kappa B mediated Up-regulation of CCCTC-binding factor in pediatric acute lymphoblastic leukemia. *Mol. Cancer* **13**, (2014).
494. Corthals, S. L. *et al.* Differential immune effects mediated by Toll-like receptors stimulation in precursor B-cell acute lymphoblastic leukaemia. *Br. J. Haematol.* **132**, 452–458 (2006).
495. Apostolov, A. *et al.* Sumoylation Inhibits the Growth Suppressive Properties of Ikaros. *PLoS One* **11**, (2016).
496. Miyai, T. *et al.* Zinc transporter SLC39A10/ZIP10 facilitates antiapoptotic signaling during early B-cell development. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 11780–11785 (2014).
497. Naqvi, A. R., Fordham, J. B. & Nares, S. miR-24, miR-30b, and miR-142-3p Regulate Phagocytosis in Myeloid Inflammatory Cells. *J. Immunol.* **194**, 1916–1927 (2015).
498. Jiffar, T. *et al.* PKC α mediates chemoresistance in acute lymphoblastic leukemia through

- effects on Bcl2 phosphorylation. *Leukemia* **18**, 505–512 (2004).
499. Buske, C. *et al.* TGF- β inhibits growth and induces apoptosis in leukemic B cell precursors. *Leukemia* **11**, 386–392 (1997).
500. Sanchez, V. E., Nichols, C., Kim, H. N., Gang, E. J. & Kim, Y. M. Targeting PI3K signaling in acute lymphoblastic leukemia. *Int. J. Mol. Sci.* **20**, (2019).
501. Zhang, S., Fritz, N., Ibarra, C. & Uhlén, P. Inositol 1,4,5-trisphosphate receptor subtype-specific regulation of calcium oscillations. *Neurochem. Res.* **36**, 1175–1185 (2011).
502. Shojaee, S. *et al.* Erk Negative Feedback Control Enables Pre-B Cell Transformation and Represents a Therapeutic Target in Acute Lymphoblastic Leukemia. *Cancer Cell* **28**, 114–128 (2015).
503. Lee, J. H. S., Vo, T. T. & Fruman, D. A. Targeting mTOR for the treatment of B cell malignancies. *Br. J. Clin. Pharmacol.* 1213–1228 (2016) doi:10.1111/bcp.12888.
504. Chen, D. W. C., Saha, V., Liu, J. Z., Schwartz, J. M. & Krstic-Demonacos, M. Erg and AP-1 as determinants of glucocorticoid response in acute lymphoblastic leukemia. *Oncogene* **32**, 3039–3048 (2013).
505. Esparza, S. D. & Sakamoto, K. M. Topics in pediatric leukemia - Acute lymphoblastic leukemia. *MedGenMed Medscape Gen. Med.* **7**, (2005).
506. Schultz, K. R. *et al.* Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: A combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children’s Cancer Group (CCG). *Blood* **109**, 926–935 (2007).

507. Ma, H., Sun, H. & Sun, X. Survival improvement by decade of patients aged 0-14 years with acute lymphoblastic leukemia: A SEER analysis. *Sci. Rep.* **4**, (2014).
508. Pulte, D., Gondos, A. & Brenner, H. Improvement in survival in younger patients with acute lymphoblastic leukemia from the 1980s to the early 21st century. *Blood* **113**, 1408–1411 (2009).
509. Kantarjian, H. *et al.* Long-term follow-up results of hyperfractionated cyclophosphamide, vincristine, doxorubicin, and dexamethasone (Hyper-CVAD), a dose-intensive regimen, in adult acute lymphocytic leukemia. *Cancer* **101**, 2788–2801 (2004).
510. Pulte, D. *et al.* Survival of adults with acute lymphoblastic leukemia in Germany and the United States. *PLoS One* **9**, (2014).
511. Sive, J. I. *et al.* Outcomes in older adults with acute lymphoblastic leukaemia (ALL): Results from the international MRC UKALL XII/ECOG2993 trial. *Br. J. Haematol.* **157**, 463–471 (2012).
512. Kelly, M. E. *et al.* Treatment of relapsed precursor-b acute lymphoblastic leukemia with intensive chemotherapy: Pog (pediatric oncology group) study 9411 (SIMAL 9). *J. Pediatr. Hematol. Oncol.* **35**, 509–513 (2013).
513. Brüggemann, M. & Kotrova, M. Minimal residual disease in adult ALL: Technical aspects and implications for correct clinical interpretation. *Hematology* (2017) doi:10.1182/asheducation-2017.1.13.
514. Bassan, R. *et al.* A systematic literature review and metaanalysis of minimal residual disease as a prognostic indicator in adult B-cell acute lymphoblastic leukemia.

- Haematologica* **104**, 2028–2039 (2019).
515. Berry, D. A. *et al.* Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: A meta-analysis. *JAMA Oncol.* **3**, (2017).
516. Rau, R. E. *et al.* Prognostic impact of minimal residual disease at the end of consolidation in NCI standard-risk B-lymphoblastic leukemia: A report from the Children’s Oncology Group. *Pediatr. Blood Cancer* **68**, (2021).
517. Tricarico, C., Clancy, J. & D’Souza-Schorey, C. Biology and biogenesis of shed microvesicles. *Small GTPases* (2017) doi:10.1080/21541248.2016.1215283.
518. Díaz-Varela, M. *et al.* Proteomics study of human cord blood reticulocyte-derived exosomes. *Sci. Rep.* (2018) doi:10.1038/s41598-018-32386-2.
519. Anand, S., Samuel, M., Kumar, S. & Mathivanan, S. Ticket to a bubble ride: Cargo sorting into exosomes and extracellular vesicles. *Biochimica et Biophysica Acta - Proteins and Proteomics* vol. 1867 (2019).
520. Longjohn, M. N. *et al.* Deciphering the messages carried by extracellular vesicles in hematological malignancies. *Blood Rev.* (2020) doi:10.1016/j.blre.2020.100734.
521. Hunter, M. P. *et al.* Detection of microRNA expression in human peripheral blood microvesicles. *PLoS One* **3**, e3694 (2008).
522. Huang, X. *et al.* Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genomics* **14**, (2013).
523. Nolte-’t Hoen, E. N. M. *et al.* Deep sequencing of RNA from immune cell-derived

- vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res.* **40**, 9272–9285 (2012).
524. Guescini, M., Genedani, S., Stocchi, V. & Agnati, L. F. Astrocytes and Glioblastoma cells release exosomes carrying mtDNA. *J. Neural Transm.* (2010) doi:10.1007/s00702-009-0288-8.
525. Jeppesen, D. K. *et al.* Reassessment of Exosome Composition. *Cell* **177**, 428-445.e18 (2019).
526. Roy, J. W. *et al.* Small RNA sequencing analysis of peptide-affinity isolated plasma extracellular vesicles distinguishes pancreatic cancer patients from non-affected individuals. *Sci. Rep.* **13**, (2023).
527. Gonda, A. *et al.* Extracellular Vesicle Molecular Signatures Characterize Metastatic Dynamicity in Ovarian Cancer. *Front. Oncol.* **11**, (2021).
528. He, X. *et al.* Circulating exosomal mRNA signatures for the early diagnosis of clear cell renal cell carcinoma. *BMC Med.* **20**, (2022).
529. Conley, A. *et al.* High-throughput sequencing of two populations of extracellular vesicles provides an mRNA signature that can be detected in the circulation of breast cancer patients. *RNA Biol.* **14**, 305–316 (2017).
530. Li, L. *et al.* The potential role of extracellular vesicle-derived small RNAs in AML research as non-invasive biomarker. A24.2-A25 (2022) doi:10.1136/jitc-2022-itoc9.45.
531. Stam, J., Bartel, S., Bischoff, R. & Wolters, J. C. Isolation of extracellular vesicles with combined enrichment methods. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1169**,

- (2021).
532. Sork, H. *et al.* Profiling of extracellular small rnas highlights a strong bias towards non-vesicular secretion. *Cells* **10**, (2021).
 533. Albanese, M. *et al.* MicroRNAs are minor constituents of extracellular vesicles that are rarely delivered to target cells. *PLoS Genet.* **17**, (2021).
 534. Balaj, L. *et al.* Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nat. Commun.* (2011) doi:10.1038/ncomms1180.
 535. Agarwal, K. *et al.* Analysis of exosome release as a cellular response to MAPK pathway inhibition. *Langmuir* (2015) doi:10.1021/acs.langmuir.5b00095.
 536. M., L. *et al.* High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients. *PLoS One* **4**, (2009).
 537. Silva, J. *et al.* Analysis of exosome release and its prognostic value in human colorectal cancer. *Genes Chromosom. Cancer* **51**, 409–418 (2012).
 538. Miljkovic-Licina, M., Arraud, N., Zahra, A. D., Ropraz, P. & Matthes, T. Quantification and phenotypic characterization of extracellular vesicles from patients with acute myeloid and B-cell lymphoblastic leukemia. *Cancers (Basel)*. **14**, (2022).
 539. Hu, W. *et al.* Comprehensive landscape of extracellular vesicle-derived RNAs in cancer initiation, progression, metastasis and cancer immunology. *Mol. Cancer* **19**, (2020).
 540. Jabalee, J., Towle, R. & Garnis, C. The Role of Extracellular Vesicles in Cancer: Cargo, Function, and Therapeutic Implications. *Cells* **7**, (2018).
 541. Clancy, J. W., Zhang, Y., Sheehan, C. & D'Souza-Schorey, C. An ARF6–Exportin-5 axis

- delivers pre-miRNA cargo to tumour microvesicles. *Nat. Cell Biol.* **21**, 856–866 (2019).
542. Van Balkom, B. W. M., Eisele, A. S., Michiel Pegtel, D., Bervoets, S. & Verhaar, M. C. Quantitative and qualitative analysis of small RNAs in human endothelial cells and exosomes provides insights into localized RNA processing, degradation and sorting. *J. Extracell. Vesicles* **4**, 1–14 (2015).
543. Chakraborty, S. K., Prakash, A., Nechooshtan, G., Hearn, S. & Gingeras, T. R. Extracellular vesicle-mediated transfer of processed and functional RNY5 RNA. *Rna* **21**, 1966–1979 (2015).
544. Kwasnik, M. *et al.* Protein-Coding Region Derived Small RNA in Exosomes from Influenza A Virus–Infected Cells. *Int. J. Mol. Sci.* **24**, (2023).
545. Xia, J., Wang, M., Zhu, Y., Bu, C. & Li, T. Differential mRNA and long noncoding RNA expression profiles in pediatric B-cell acute lymphoblastic leukemia patients. *BMC Pediatr.* **22**, (2022).
546. Longjohn, M. N., Squires, W. R. B. & Christian, S. L. Meta-analysis of microRNA profiling data does not reveal a consensus signature for B cell acute lymphoblastic leukemia. *Gene* **821**, (2022).
547. Verbeek, M. W. C., Erkeland, S. J. & van der Velden, V. H. J. Dysregulation of Small Nucleolar RNAs in B-Cell Malignancies. *Biomedicines* **10**, (2022).
548. Zhu, Y. *et al.* Mechanism of circADD2 as ceRNA in Childhood Acute Lymphoblastic Leukemia. *Front. Cell Dev. Biol.* **9**, (2021).
549. Zhan, T. *et al.* Mir-455-3p functions as a tumor suppressor by restraining wnt/ β -catenin

- signaling via taz in pancreatic cancer. *Cancer Manag. Res.* **12**, 1483–1492 (2020).
550. Yi, X., Wang, Y. & Xu, S. MiR-455-3p downregulation facilitates cell proliferation and invasion and predicts poor prognosis of osteosarcoma. *J. Orthop. Surg. Res.* **15**, (2020).
551. Gao, X. *et al.* miR-455-3p serves as prognostic factor and regulates the proliferation and migration of non-small cell lung cancer through targeting HOXB5. *Biochem. Biophys. Res. Commun.* **495**, 1074–1080 (2018).
552. Márton, É. *et al.* Comparative analysis of cell-free mir-205-5p, let-7f-5p, and mir-483-5p expression in ovarian cell cultures and plasma samples of patients with ovarian cancer. *Appl. Sci.* **11**, 1–10 (2021).
553. Niu, M. *et al.* High expression of miR-25 predicts favorable chemotherapy outcome in patients with acute myeloid leukemia. *Cancer Cell Int.* **19**, (2019).
554. Xiong, Q. *et al.* Characterization of miRNomes in Acute and Chronic Myeloid Leukemia Cell Lines. *Genomics, Proteomics Bioinforma.* **12**, 79–91 (2014).
555. Rzepiel, A. *et al.* MiR-128-3p as blood based liquid biopsy biomarker in childhood acute lymphoblastic leukemia. *Mol. Cell. Probes* **67**, (2023).
556. Mets, E. *et al.* MicroRNA-128-3p is a novel oncomiR targeting PHF6 in T-cell acute lymphoblastic leukemia. *Haematologica* **99**, 1326–1333 (2014).
557. Martins, J. R. B. *et al.* MiR-125a-3p and miR-320b differentially expressed in patients with chronic myeloid leukemia treated with allogeneic hematopoietic stem cell transplantation and imatinib mesylate. *Int. J. Mol. Sci.* **22**, (2021).
558. Blume, C. J. *et al.* P53-dependent non-coding RNA networks in chronic lymphocytic

- leukemia. *Leukemia* **29**, 2015–2023 (2015).
559. Ma, L., Yang, H. & Yang, X. Identification and integrative analysis of microRNAs in myelodysplastic syndromes based on microRNAs expression profile . *Precis. Med. Sci.* **10**, 142–150 (2021).
560. Nemes, K. *et al.* Expression of Certain Leukemia/Lymphoma Related microRNAs and its Correlation with Prognosis in Childhood Acute Lymphoblastic Leukemia. *Pathol. Oncol. Res.* **21**, 597–604 (2015).
561. Zhang, H. *et al.* MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia. *PLoS One* **4**, (2009).
562. Organista-Nava, J. *et al.* High miR-24 expression is associated with risk of relapse and poor survival in acute leukemia. *Oncol. Rep.* **33**, 1639–1649 (2015).
563. Yan, W. *et al.* Extracellular vesicles carrying miRNA-181b-5p affects the malignant progression of acute lymphoblastic leukemia. *J. Transl. Med.* **19**, (2021).
564. Moussay, E. *et al.* MicroRNA as biomarkers and regulators in B-cell chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6573–6578 (2011).
565. Ninawe, A. *et al.* MiR-486-5p: A Prognostic Biomarker for Chronic Myeloid Leukemia. *ACS Omega* **6**, 7711–7718 (2021).
566. Gregory, G. L. & Cople, I. M. Modulating the expression of tumor suppressor genes using activating oligonucleotide technologies as a therapeutic approach in cancer. *Mol. Ther. - Nucleic Acids* **31**, 211–223 (2023).
567. Shen, Z. *et al.* Characterization of microRNA expression profiles by deep sequencing of

- small RNA libraries in leukemia patients from Naxi ethnic. *Transl. Cancer Res.* **8**, 160–169 (2019).
568. Garg, P., Jamal, F. & Srivastava, P. Deciphering the role of precursor miR-12136 and miR-8485 in the progression of intellectual disability (ID). *IBRO Neurosci. Reports* **13**, 393–401 (2022).
569. Lu, S., Chen, Z., Liu, Z. & Liu, Z. Unmasking the biological function and regulatory mechanism of NOC2L: a novel inhibitor of histone acetyltransferase. *J. Transl. Med.* **21**, (2023).
570. Cho, Y. *et al.* Per1 enhances mitochondrial biogenesis, oxidative capacity, and fatigue resistance in adult skeletal muscle. *FASEB J.* **30**, 674–687 (2016).
571. Wang, C., Chen, Y., Chen, K. & Zhang, L. Long Noncoding RNA LINC01134 Promotes Hepatocellular Carcinoma Metastasis via Activating AKT1S1 and NF- κ B Signaling. *Front. Cell Dev. Biol.* **8**, (2020).
572. Shen, S. *et al.* LINC01714 Enhances Gemcitabine Sensitivity by Modulating FOXO3 Phosphorylation in Cholangiocarcinoma. *Mol. Ther. - Nucleic Acids* **19**, 446–457 (2020).
573. Luo, S. *et al.* How Does Complement Affect Hematological Malignancies: From Basic Mechanisms to Clinical Application. *Front. Immunol.* **11**, (2020).
574. Zapata-García, J. A. *et al.* Comparative Genomic Hybridization and Transcriptome Sequencing Reveal Genes with Gain in Acute Lymphoblastic Leukemia: JUP Expression Emerges as a Survival-Related Gene. *Diagnostics* **12**, (2022).
575. Zheng, X. *et al.* Identification of a seven-lncRNAs panel that serves as a prognosis

- predictor and contributes to the malignant progression of laryngeal squamous cell carcinoma. *Front. Oncol.* **13**, (2023).
576. Min, S. *et al.* MRPS31 loss is a key driver of mitochondrial deregulation and hepatocellular carcinoma aggressiveness. *Cell Death Dis.* **12**, (2021).
577. Zheng, J. *et al.* Integrative Analysis of Multi-Omics Identified the Prognostic Biomarkers in Acute Myelogenous Leukemia. *Front. Oncol.* **10**, (2020).
578. Lee, R. *et al.* PDE7A is expressed in human B-lymphocytes and is up-regulated by elevation of intracellular cAMP. *Cell. Signal.* **14**, 277–284 (2002).
579. Rehn, J. A., O’connor, M. J., White, D. L. & Yeung, D. T. DUX hunting—clinical features and diagnostic challenges associated with DUX4-rearranged leukaemia. *Cancers (Basel)*. **12**, 1–15 (2020).
580. Russel, L., Enshaei, A., Jones, L. & Al. IGH@ translocations are prevalent in teenagers and young adults with acute lymphoblastic leukemia and are associated with a poor outcome. *J Clin Oncol* (2014).
581. Fu, D. *et al.* Prognosis and Characterization of Immune Microenvironment in Acute Myeloid Leukemia Through Identification of an Autophagy-Related Signature. *Front. Immunol.* **12**, (2021).
582. Roy Choudhury, S. *et al.* Epigenetically Enhanced MED12L in ETO2-GLIS2 Positive Pediatric Acute Megakaryoblastic Leukemia Is Associated with Resistance to the CDK8 Inhibitors. *Blood* **138**, 2208–2208 (2021).
583. Qu, K. *et al.* MCM7 promotes cancer progression through cyclin D1-dependent signaling

- and serves as a prognostic marker for patients with hepatocellular carcinoma. *Cell Death Dis.* **8**, (2017).
584. Zhang, S. & Sun, Y. Targeting CDC34 E2 ubiquitin conjugating enzyme for lung cancer therapy. *EBioMedicine* **54**, (2020).
585. Jiang, F., Huang, X., Yang, X., Zhou, H. & Wang, Y. NUF2 Expression Promotes Lung Adenocarcinoma Progression and Is Associated With Poor Prognosis. *Front. Oncol.* **12**, (2022).
586. Castro, A. *et al.* Elevated neoantigen levels in tumors with somatic mutations in the HLA-A, HLA-B, HLA-C and B2M genes. *BMC Med. Genomics* **12**, (2019).
587. MOHAMED EL-SHANSHORY, M.D., E. E.-S. M. S. ., LAILA M. SHREIF, M.D., S. A. M. D. . & ENAS A.A. BAKI, Ph.D., A.-A. A. Z. P. D. . miRNA Expression Profiling in Pediatric B-Cell Acute Lymphoblastic Leukemia by Microarray Technology. *Med. J. Cairo Univ.* **86**, 1049–1053 (2018).
588. Hadavi, R. *et al.* Expression of Bioinformatically Candidate miRNAs including, miR-576-5p, miR-501-3p and miR-3143, Targeting PI3K Pathway in Triple-Negative Breast Cancer. *Galen Med. J.* **8**, 1646 (2019).
589. Lv, M. *et al.* B-cell acute lymphoblastic leukemia-related microRNAs: uncovering their diverse and special roles. *Am. J. Cancer Res.* **11**, 1104–1120 (2021).
590. Valiollahi, E., Ribera, J. M., Genescà, E. & Behravan, J. Genome-wide identification of microRNA signatures associated with stem/progenitor cells in Philadelphia chromosome-positive acute lymphoblastic leukemia. *Mol. Biol. Rep.* **46**, 1295–1306 (2019).

591. Poel, D. *et al.* Expression of let-7i and miR-192 is associated with resistance to cisplatin-based chemoradiotherapy in patients with larynx and hypopharynx cancer. *Oral Oncol.* **109**, (2020).
592. Zhou, H. *et al.* Downregulation of miR-224 and let-7i contribute to cell survival and chemoresistance in chronic myeloid leukemia cells by regulating ST3GAL IV expression. *Gene* **626**, 106–118 (2017).
593. Ransohoff, J. D., Wei, Y. & Khavari, P. A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **19**, 143–157 (2018).
594. Hutarew, G. *et al.* Methylome Profiling of PD-L1-Expressing Glioblastomas Shows Enrichment of Post-Transcriptional and RNA-Associated Gene Regulation. *Cancers (Basel)*. **14**, (2022).
595. He, Y., Ye, Y., Tian, W. & Qiu, H. A Novel lncRNA Panel Related to Ferroptosis, Tumor Progression, and Microenvironment is a Robust Prognostic Indicator for Glioma Patients. *Front. Cell Dev. Biol.* **9**, (2021).
596. Pasięka, R., Zasoński, G. & Raczyńska, K. D. Role of Long Intergenic Noncoding RNAs in Cancers with an Overview of MicroRNA Binding. *Mol. Diagnosis Ther.* **27**, 29–47 (2023).
597. Zhou, C. *et al.* Comprehensive analysis of GINS subunits prognostic value and ceRNA network in sarcoma. *Front. Cell Dev. Biol.* **10**, (2022).
598. Sun, Q. *et al.* Identification and Validation of 17-lncRNA Related to Regulatory T Cell Heterogeneity as a Prognostic Signature for Head and Neck Squamous Cell Carcinoma.

- Front. Immunol.* **12**, (2021).
599. Peixoto, C. *et al.* Identification of biomarkers predictive of metastasis development in early-stage colorectal cancer using network-based regularization. *BMC Bioinformatics* **24**, (2023).
600. Sánchez-Beato, M., Sánchez-Aguilera, A. & Piris, M. A. Cell cycle deregulation in B-cell lymphomas. *Blood* **101**, 1220–1235 (2003).
601. Huang, M. & Zhu, J. The Regulation of Normal and Leukemic Hematopoietic Stem Cells by Niches. *Cancer Microenviron.* **5**, 295–305 (2012).
602. Filippi, M. D. & Ghaffari, S. Mitochondria in the maintenance of hematopoietic stem cells: New perspectives and opportunities. *Blood* **133**, 1943–1952 (2019).
603. Papa, L., Djedaini, M. & Hoffman, R. Mitochondrial role in stemness and differentiation of hematopoietic stem cells. *Stem Cells Int.* **2019**, (2019).
604. Lang, F., Wojcik, B. & Rieger, M. A. Stem Cell Hierarchy and Clonal Evolution in Acute Lymphoblastic Leukemia. *Stem Cells Int.* **2015**, (2015).
605. Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327–334 (2014).
606. Chen, Y., Li, J. & Zhao, Z. Redox control in acute lymphoblastic leukemia: From physiology to pathology and therapeutic opportunities. *Cells* **10**, (2021).
607. Ghaffari, S. Oxidative stress in the regulation of normal and neoplastic hematopoiesis. *Antioxidants Redox Signal.* **10**, 1923–1940 (2008).
608. Ludin, A. *et al.* Reactive oxygen species regulate hematopoietic stem cell self-renewal,

- migration and development, as well as their bone marrow microenvironment. *Antioxidants Redox Signal.* **21**, 1605–1619 (2014).
609. Ghelli Luserna Di Rora, A., Iacobucci, I. & Martinelli, G. The cell cycle checkpoint inhibitors in the treatment of leukemias. *J. Hematol. Oncol.* **10**, (2017).
610. Gururaja Rao, S., Patel, N. J. & Singh, H. Intracellular Chloride Channels: Novel Biomarkers in Diseases. *Front. Physiol.* **11**, (2020).
611. Tompkins, V. S., Rouse, W. B., O’Leary, C. A., Andrews, R. J. & Moss, W. N. Analyses of human cancer driver genes uncovers evolutionarily conserved RNA structural elements involved in posttranscriptional control. *PLoS One* **17**, (2022).
612. Ueno, H. *et al.* Landscape of driver mutations and their clinical impacts in pediatric B-cell precursor acute lymphoblastic leukemia. *Blood Adv.* **4**, 5165–5173 (2020).
613. Johnson, S. M. *et al.* Acute lymphoblastic leukaemia cells produce large extracellular vesicles containing organelles and an active cytoskeleton. *J. Extracell. Vesicles* (2017) doi:10.1080/20013078.2017.1294339.
614. Farahani, M., Rubbi, C., Liu, L., Slupsky, J. R. & Kalakonda, N. CLL exosomes modulate the transcriptome and behaviour of recipient stromal cells and are selectively enriched in MIR-202-3p. *PLoS One* (2015) doi:10.1371/journal.pone.0141429.
615. Mullighan, C. G. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am. Soc. Hematol. Educ. Program* **2012**, 389–396 (2012).
616. Wu, H., Wang, C. & Wu, Z. PROPER: Comprehensive power evaluation for differential expression using RNA-seq. *Bioinformatics* **31**, 233–241 (2015).

617. The Leukemia & Lymphoma Society of Canada. Blood cancer in Canada. Facts and Stats. *Leukemia and Lymphoma Society of Canada*
http://www.llscanada.org/sites/default/files/National/CANADA/Pdf/InfoBooklets/Blood_Cancer_in_Canada_Facts_%26_Stats_2016.pdf (2016).
618. Momen-Heravi, F. & Bala, S. Extracellular vesicles in oral squamous carcinoma carry oncogenic miRNA profile and reprogram monocytes via NF- κ B pathway. *Oncotarget* (2018) doi:10.18632/oncotarget.26208.
619. Al-Nedawi, K. *et al.* Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat. Cell Biol.* (2008) doi:10.1038/ncb1725.
620. Patel, N. *et al.* A dyad of lymphoblastic lysosomal cysteine proteases degrades the antileukemic drug L-asparaginase. *J. Clin. Invest.* (2009) doi:10.1172/JCI37977.
621. Lorincz, Á. M. *et al.* Effect of storage on physical and functional properties of extracellular vesicles derived from neutrophilic granulocytes. *J. Extracell. Vesicles* (2014) doi:10.3402/jev.v3.25465.
622. Matsumoto, J., Stewart, T., Banks, W. A. & Zhang, J. The Transport Mechanism of Extracellular Vesicles at the Blood-Brain Barrier. *Curr. Pharm. Des.* (2018) doi:10.2174/1381612823666170913164738.
623. Gourlay, J. *et al.* The emergent role of exosomes in glioma. *J. Clin. Neurosci.* **35**, 13–23 (2017).
624. Thakur, B. K. *et al.* Detection of AML-specific mutations in pediatric patient plasma using extracellular vesicle-derived RNA. *Ann. Hematol.* (2019) doi:10.1007/s00277-019-

03608-y.

625. Nunes, P. C. (University of Porto, P. & Caires, H. R., Sobrinho-Simoes, M. A., Vasconcelow, M. H. Circulating EVs for AML minimal residual disease biomarkers detection. *Porto Biomed. J.* **6**, 5–9 (2017).
626. Yoshida, Y., Kobune, M., Miura, S. & Masahiro Yoshida, Masayoshi Kobune, Shogo Miura, Soushi Ibata, Satoshi Iyama, Tsutomu Sato, Kazuyuki Murase, Kohichi Takada, Kaoru Ono, Akari Hashimoto, Ayumi Tatekoshi, Yusuke Kamihara, Yusuke Sugama, Shohei Kikuchi, Hiroshi Ikeda, Hiroto Horiguchi, Yuta, M. M. and J. K. Extracellular Vesicle microRNAs from Acute Myeloid Leukemia Are Involved in the Regulation of Adherence Junction in Bone Marrow Microenvironment. *Blood J.* **128**, 2863 (2016).
627. KIM, S. J. *et al.* The Potential of Exosomes Derived from Chronic Myelogenous Leukaemia Cells as a Biomarker. *Anticancer Res.* (2018) doi:10.21873/anticancer.12679.
628. Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **9**, 654–659 (2007).
629. Cheng, L., Sharples, R. A., Scicluna, B. J. & Hill, A. F. Exosomes provide a protective and enriched source of miRNA for biomarker profiling compared to intracellular and cell-free blood. *J. Extracell. Vesicles* **3**, (2014).
630. Yeri, A. *et al.* Total extracellular small RNA profiles from plasma, saliva, and urine of healthy subjects. *Sci. Rep.* **7**, (2017).
631. Vu, L. T., Gong, J., Pham, T. T., Kim, Y. & Le, M. T. N. microRNA exchange via extracellular vesicles in cancer. *Cell Prolif.* **53**, (2020).

632. O'Grady, T. *et al.* Sorting and packaging of RNA into extracellular vesicles shape intracellular transcript levels. *BMC Biol.* **20**, (2022).
633. Hinger, S. A. *et al.* Diverse Long RNAs Are Differentially Sorted into Extracellular Vesicles Secreted by Colorectal Cancer Cells. *Cell Rep.* **25**, 715-725.e4 (2018).
634. Ahadi, A., Brennan, S., Kennedy, P. J., Hutvagner, G. & Tran, N. Long non-coding RNAs harboring miRNA seed regions are enriched in prostate cancer exosomes. *Sci. Rep.* **6**, (2016).
635. Pérez-Boza, J., Lion, M. & Struman, I. Exploring the RNA landscape of endothelial exosomes. *Rna* **24**, 423–435 (2018).
636. Batagov, A. O. & Kurochkin, I. V. Exosomes secreted by human cells transport largely mRNA fragments that are enriched in the 3'-untranslated regions. *Biol. Direct* **8**, (2013).
637. Morhayim, J. *et al.* Molecular characterization of human osteoblast-derived extracellular vesicle mRNA using next-generation sequencing. *Biochim. Biophys. Acta - Mol. Cell Res.* **1864**, 1133–1141 (2017).
638. Wei, Z. *et al.* Coding and noncoding landscape of extracellular RNA released by human glioma stem cells. *Nat. Commun.* **8**, (2017).
639. Li, Y. *et al.* Extracellular vesicles long RNA sequencing reveals abundant mRNA, circRNA, and lncRNA in human blood as potential biomarkers for cancer diagnosis. *Clin. Chem.* **65**, 798–808 (2019).
640. Weng, Q. *et al.* Extracellular vesicles-associated tRNA-derived fragments (tRFs): biogenesis, biological functions, and their role as potential biomarkers in human diseases.

- J. Mol. Med.* **100**, 679–695 (2022).
641. Gámbaro, F. *et al.* Stable tRNA halves can be sorted into extracellular vesicles and delivered to recipient cells in a concentration-dependent manner. *RNA Biol.* **17**, 1168–1182 (2020).
642. Fitz, N. F., Wang, J., Kamboh, M. I., Koldamova, R. & Lefterov, I. Small nucleolar RNAs in plasma extracellular vesicles and their discriminatory power as diagnostic biomarkers of Alzheimer’s disease. *Neurobiol. Dis.* **159**, (2021).
643. Rai, A. K. *et al.* Spaceflight-Associated Changes of snoRNAs in Peripheral Blood Mononuclear Cells and Plasma Exosomes—A Pilot Study. *Front. Cardiovasc. Med.* **9**, (2022).
644. Lunavat, T. R. *et al.* Small RNA deep sequencing discriminates subsets of extracellular vesicles released by melanoma cells – Evidence of unique microRNA cargos. *RNA Biol.* **12**, 810–823 (2015).
645. Tosar, J. P. *et al.* Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. *Nucleic Acids Res.* (2015) doi:10.1093/nar/gkv432.
646. Hobor, F. *et al.* A cryptic RNA-binding domain mediates Syncrip recognition and exosomal partitioning of miRNA targets. *Nat. Commun.* **9**, (2018).
647. Villarroya-Beltri, C. *et al.* Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nat. Commun.* **4**, 2980 (2013).
648. Santangelo, L. *et al.* The RNA-Binding Protein SYNCRIP Is a Component of the

- Hepatocyte Exosomal Machinery Controlling MicroRNA Sorting. *Cell Rep.* (2016)
doi:10.1016/j.celrep.2016.09.031.
649. Mukherjee, K. *et al.* Reversible HuR-micro RNA binding controls extracellular export of miR-122 and augments stress response. *EMBO Rep.* **17**, 1184–1203 (2016).
650. Teng, Y. *et al.* MVP-mediated exosomal sorting of miR-193a promotes colon cancer progression. *Nat. Commun.* **8**, (2017).
651. Shurtleff, M. J. *et al.* Broad role for YBX1 in defining the small noncoding RNA composition of exosomes. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E8987–E8995 (2017).
652. Zhang, J. *et al.* Exosome and exosomal microRNA: Trafficking, sorting, and function. *Genomics, Proteomics Bioinforma.* **13**, 17–24 (2015).
653. Koppers-Lalic, D. *et al.* Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. *Cell Rep.* **8**, 1649–1658 (2014).
654. Kosaka, N. *et al.* Secretory mechanisms and intercellular transfer of microRNAs in living cells. *J. Biol. Chem.* **285**, 17442–17452 (2010).
655. Mittelbrunn, M. *et al.* Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. *Nat. Commun.* **2**, (2011).
656. Groot, M. & Lee, H. Sorting Mechanisms for MicroRNAs into Extracellular Vesicles and Their Associated Diseases. *Cells* **9**, (2020).
657. Fabbiano, F. *et al.* RNA packaging into extracellular vesicles: An orchestra of RNA-binding proteins? *J. Extracell. Vesicles* **10**, (2020).
658. Chiou, N. T., Kageyama, R. & Ansel, K. M. Selective Export into Extracellular Vesicles

- and Function of tRNA Fragments during T Cell Activation. *Cell Rep.* **25**, 3356-3370.e4 (2018).
659. U., L. *et al.* Small RNAs derived from tRNAs and rRNAs are highly enriched in exosomes from both old and new world *Leishmania* providing evidence for conserved exosomal RNA Packaging. *BMC Genomics* **16**, 151 (2015).
660. Zhao, F. *et al.* Characterization of serum small extracellular vesicles and their small RNA contents across humans, rats, and mice. *Sci. Rep.* **10**, (2020).
661. Quek, C. *et al.* Defining the purity of exosomes required for diagnostic profiling of small RNA suitable for biomarker discovery. *RNA Biol.* **14**, 245–258 (2017).
662. Chen, J., Xu, Q., Zhang, Y. & Zhang, H. RNA Profiling Analysis of the Serum Exosomes Derived from Patients with Chronic Hepatitis and Acute-on-chronic Liver Failure Caused By HBV. *Sci. Rep.* **10**, (2020).
663. Chettimada, S., Lorenz, D. R., Misra, V., Wolinsky, S. M. & Gabuzda, D. Small RNA sequencing of extracellular vesicles identifies circulating miRNAs related to inflammation and oxidative stress in HIV patients. *BMC Immunol.* **21**, (2020).
664. Zheng, D. *et al.* The Role of Exosomes and Exosomal MicroRNA in Cardiovascular Disease. *Front. Cell Dev. Biol.* **8**, (2021).
665. Chevillet, J. R. *et al.* Quantitative and stoichiometric analysis of the microRNA content of exosomes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 14888–14893 (2014).
666. Costanzi, E., Romani, R., Scarpelli, P. & Bellezza, I. Extracellular vesicles-mediated transfer of miRNA LET-7b from PC3 cells to macrophages. *Genes (Basel)*. **11**, 1–10

- (2020).
667. Melo, S. A. *et al.* Cancer Exosomes Perform Cell-Independent MicroRNA Biogenesis and Promote Tumorigenesis. *Cancer Cell* **26**, 707–721 (2014).
 668. Squadrito, M. L. *et al.* Endogenous RNAs Modulate MicroRNA Sorting to Exosomes and Transfer to Acceptor Cells. *Cell Rep.* **8**, 1432–1446 (2014).
 669. Chen, T. S. & Lim, S. K. Measurement of precursor miRNA in exosomes from human ESC-derived mesenchymal stem cells. *Methods Mol. Biol.* **1024**, 69–86 (2013).
 670. Horibe, S., Tanahashi, T., Kawauchi, S., Murakami, Y. & Rikitake, Y. Mechanism of recipient cell-dependent differences in exosome uptake. *BMC Cancer* **18**, (2018).
 671. Taylor, D. D. & Gercel-Taylor, C. MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. *Gynecol. Oncol.* **110**, 13–21 (2008).
 672. Kloecker, G. H., Rabinowits, G., Gercel-Taylor, C., Day, J. M. & Taylor, D. D. Exosomal microRNA: A Diagnostic Marker for Lung Cancer. *Clin. Lung Cancer* **9**, 295 (2008).
 673. Skog, J. *et al.* Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* **10**, 1470–1476 (2008).
 674. Luo, X., Jean-Toussaint, R., Sacan, A. & Ajit, S. K. Differential RNA packaging into small extracellular vesicles by neurons and astrocytes. *Cell Commun. Signal.* **19**, (2021).
 675. Chan, J. J., Tabatabaeian, H. & Tay, Y. 3'UTR heterogeneity and cancer progression. *Trends Cell Biol.* (2022) doi:10.1016/j.tcb.2022.10.001.
 676. Koh, W. S., Porter, J. R. & Batchelor, E. Tuning of mRNA stability through altering 3'-UTR sequences generates distinct output expression in a synthetic circuit driven by p53

- oscillations. *Sci. Rep.* **9**, (2019).
677. Khabar, K. S. A. Hallmarks of cancer and AU-rich elements. *Wiley Interdiscip. Rev. RNA* **8**, (2017).
678. Beauregard, A. P. *et al.* Pax-5 protein expression is regulated by transcriptional 3'utr editing. *Cells* **11**, (2022).
679. Velázquez-Avila, M. *et al.* High cortactin expression in B-cell acute lymphoblastic leukemia is associated with increased transendothelial migration and bone marrow relapse. *Leukemia* **33**, 1337–1348 (2019).
680. Seal, R. L. *et al.* A guide to naming human non-coding RNA genes. *EMBO J.* **39**, (2020).
681. James, A. R. *et al.* Long non-coding RNAs defining major subtypes of B cell precursor acute lymphoblastic leukemia. *J. Hematol. Oncol.* **12**, (2019).
682. T.R., F. *et al.* LncRNA expression discriminates karyotype and predicts survival in B-lymphoblastic leukemia. *Mol. Cancer Res.* **13**, 839–851 (2015).
683. Lajoie, M. *et al.* Specific expression of novel long non-coding RNAs in high-hyperdiploid childhood acute lymphoblastic leukemia. *PLoS One* **12**, (2017).
684. Cuadros, M. *et al.* Expression of the long non-coding RNA TCL6 is associated with clinical outcome in pediatric B-cell acute lymphoblastic leukemia. *Blood Cancer J.* **9**, (2019).
685. Bárcenas-López, D. A. *et al.* Transcriptome analysis identifies LINC00152 as a biomarker of early relapse and mortality in acute lymphoblastic leukemia. *Genes (Basel)*. **11**, (2020).
686. Affinito, O. *et al.* LncRNAs–mRNAs co–expression network underlying childhood B–cell

- acute lymphoblastic leukaemia: A pilot study. *Cancers (Basel)*. **12**, 1–20 (2020).
687. Liu, Y. *et al.* Knockdown of ADORA2A antisense RNA 1 inhibits cell proliferation and enhances imatinib sensitivity in chronic myeloid leukemia. *Bioengineered* **13**, 2296–2307 (2022).
688. Rodríguez-Malavé, N. I. *et al.* BALR-6 regulates cell growth and cell survival in B-lymphoblastic leukemia. *Mol. Cancer* **14**, (2015).
689. Liu, Z. *et al.* Mathematical models of amino acid panel for assisting diagnosis of children acute leukemia. *J. Transl. Med.* **17**, (2019).
690. Ferguson, D. *et al.* Amino Acid Stress Response Genes Promote L-Asparaginase Resistance in Pediatric Acute Lymphoblastic Leukemia. *Blood* **138**, 3304–3304 (2021).
691. Sun, C. *et al.* tRNA-Derived Fragments as Novel Predictive Biomarkers for Trastuzumab-Resistant Breast Cancer. *Cell. Physiol. Biochem.* **49**, 419–431 (2018).
692. Bowen, R. A. R., Sattayapiwat, A., Gounden, V. & Remaley, A. T. Blood collection tube-related alterations in analyte concentrations in quality control material and serum specimens. *Clin. Biochem.* **47**, 150–157 (2014).
693. Lima-Oliveira, G. *et al.* Sodium citrate vacuum tubes validation: Preventing preanalytical variability in routine coagulation testing. *Blood Coagul. Fibrinolysis* **24**, 252–255 (2013).
694. Banfi, G., Salvagno, G. L. & Lippi, G. The role of ethylenediamine tetraacetic acid (EDTA) as in vitro anticoagulant for diagnostic purposes. *Clin. Chem. Lab. Med.* **45**, 565–576 (2007).
695. Li, G. *et al.* Comparison of glucose determinations on blood samples collected in three

- types of tubes. *Ann. Clin. Lab. Sci.* **43**, 278–284 (2013).
696. Eisinger, S. W., Schwartz, M., Dam, L. & Riedel, S. Evaluation of the BD vacutainer plus urine C&S preservative tubes compared with nonpreservative urine samples stored at 4°C and room temperature. *Am. J. Clin. Pathol.* **140**, 306–313 (2013).
697. Medina Diaz, I. *et al.* Performance of Streck cfDNA blood collection tubes for liquid biopsy testing. *PLoS One* **11**, (2016).
698. Martellucci, S. *et al.* Extracellular vesicles: New endogenous shuttles for mirnas in cancer diagnosis and therapy? *Int. J. Mol. Sci.* **21**, 1–24 (2020).
699. Chiam, K. *et al.* Serum outperforms plasma in small extracellular vesicle microRNA biomarker studies of adenocarcinoma of the esophagus. *World J. Gastroenterol.* **26**, 2570–2583 (2020).
700. Wachalska, M. *et al.* Protein Complexes in Urine Interfere with Extracellular Vesicle Biomarker Studies. *J. Circ. Biomarkers* **5**, (2016).
701. Erdbrügger, U. *et al.* Urinary extracellular vesicles: A position paper by the Urine Task Force of the International Society for Extracellular Vesicles. *J. Extracell. Vesicles* **10**, (2021).
702. Fortunato, D. *et al.* Selective isolation of extracellular vesicles from minimally processed human plasma as a translational strategy for liquid biopsies. *Biomark. Res.* **10**, (2022).
703. P., W. *et al.* Circulating MicroRNAs as novel biomarkers for platelet activation. *Circ. Res.* **112**, 595–600 (2013).
704. Veryaskina, Y. A. *et al.* Selection of reference genes for quantitative analysis of

- microRNA expression in three different types of cancer. *PLoS One* **17**, (2022).
705. Zhou, S. *et al.* MicroRNA-192-5p suppresses the initiation and progression of osteosarcoma by targeting USP1. *Oncol. Lett.* **15**, 6947–6956 (2018).
706. Jain, G. *et al.* Urinary extracellular vesicles miRNA—A new era of prostate cancer biomarkers. *Front. Genet.* **14**, (2023).
707. Sun, C. *et al.* MicroRNA-181a-5p Promotes Osteosarcoma Progression via PTEN/AKT Pathway. *Anal. Cell. Pathol.* **2022**, (2022).
708. Jiang, X. *et al.* Identification and Validation Prognostic Impact of MiRNA-30a-5p in Lung Adenocarcinoma. *Front. Oncol.* **12**, (2022).
709. Sheedy, P. & Medarova, Z. The fundamental role of miR-10b in metastatic cancer. *Am. J. Cancer Res.* **8**, 1674–1688 (2018).
710. Krammer, T. L. *et al.* Comprehensive Characterization of Platelet-Enriched MicroRNAs as Biomarkers of Platelet Activation. *Cells* **11**, (2022).
711. Nagalla, S. *et al.* Platelet MicromRNA-mRNA Co-Expression Profiles Correlate with Platelet Reactivity. *Blood* **116**, 2016–2016 (2010).
712. Yamada, M., Funaki, S. & Miki, S. Formaldehyde interacts with RNA rather than DNA: Accumulation of formaldehyde by the RNA-inorganic hybrid material. *Int. J. Biol. Macromol.* **122**, 168–173 (2019).
713. Muller, L., Hong, C. S., Stolz, D. B., Watkins, S. C. & Whiteside, T. L. Isolation of biologically-active exosomes from human plasma. *J. Immunol. Methods* **411**, 55–65 (2014).

714. Helwak, A., Kudla, G., Dudnakova, T. & Tollervey, D. Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. *Cell* **153**, 654–665 (2013).
715. Kim, S. *et al.* The regulatory impact of RNA-binding proteins on microRNA targeting. *Nat. Commun.* **12**, (2021).
716. Ramanathan, M., Porter, D. F. & Khavari, P. A. Methods to study RNA–protein interactions. *Nat. Methods* **16**, 225–234 (2019).
717. Behrens, A., Rodschinka, G. & Nedialkova, D. D. High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. *Mol. Cell* **81**, 1802-1815.e7 (2021).
718. Koga, Y. *et al.* Exosome can prevent RNase from degrading microRNA in feces. *J. Gastrointest. Oncol.* **2**, 215–22 (2011).
719. Arroyo, J. D. *et al.* Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5003–5008 (2011).
720. Weber, J. A. *et al.* The microRNA spectrum in 12 body fluids. *Clin. Chem.* **56**, 1733–1741 (2010).
721. O’Brien, K., Breyne, K., Ughetto, S., Laurent, L. C. & Breakefield, X. O. RNA delivery by extracellular vesicles in mammalian cells and its applications. *Nat. Rev. Mol. Cell Biol.* **21**, 585–606 (2020).
722. Miranda, K. C. *et al.* Nucleic acids within urinary exosomes/microvesicles are potential biomarkers for renal disease. *Kidney Int.* **78**, 191–199 (2010).

723. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
724. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).
725. Munagala, R. *et al.* Exosome-mediated delivery of RNA and DNA for gene therapy. *Cancer Lett.* **505**, 58–72 (2021).
726. Ahmadi, S. E. *et al.* Viral vectors and extracellular vesicles: innate delivery systems utilized in CRISPR/Cas-mediated cancer therapy. *Cancer Gene Ther.* (2023)
doi:10.1038/s41417-023-00597-z.
727. Collison, L. W. *et al.* IL-35-mediated induction of a potent regulatory T cell population. *Nat. Immunol.* **11**, 1093–1101 (2010).
728. Sullivan, J. A. *et al.* Treg-Cell-Derived IL-35-Coated Extracellular Vesicles Promote Infectious Tolerance. *Cell Rep.* **30**, 1039-1051.e5 (2020).

Appendices

Appendix A: Copyright licenses

License for Longjohn et al, 2019 – Blood Reviews

License for Longjohn et al, 2022 – Gene

From: Thomas Rexson Yesudoss

Date: Tuesday, July 11, 2023 03:55 PM GMT

Dear Modeline Longjohn,

Thank you for contacting us.

As an Elsevier journal author, you retain the right to Include the article in a thesis or dissertation (provided that this is not to be published commercially) whether in full or in part, subject to proper acknowledgment; see <https://www.elsevier.com/about/policies/copyright> for more information.

As this is a retained right, no written permission from Elsevier is necessary.

As outlined in our permissions licenses, this extends to the posting to your university's digital repository of the thesis provided that if you include the published journal article (PJA) version, it is embedded in your thesis only and not separately downloadable.


Thank you.

Kind regards,

Thomas Rexson Yesudoss

Copyrights Specialist

ELSEVIER | HCM - Health Content Management



Deciphering the messages carried by extracellular vesicles in hematological malignancies
Author: Modeline N. Longjohn,Jo-Anna B.J. Hudson,Nicole C. Smith,Matthew L. Rise,Paul C. Moorehead,Sherril L. Christian
Publication: Blood Reviews
Publisher: Elsevier
Date: March 2021
© 2020 Elsevier Ltd. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

[BACK](#) [CLOSE WINDOW](#)

Appendix B: Codes and scripts for bioinformatics

B1: *Ethics approval for pediatric B-ALL EV small RNA signature and packaging study*

From: j.hudson.j.hudson@mun.ca
Subject: Fwd: HREB - Approval of Ethics Renewal 587782
Date: May 24, 2023 at 9:32 AM
To: Sherr Christian sherr@mum.ca

JH

This was the email confirmation of approval and then i update the fall for the next additional year

Cheers,
Jo-Anna

----- Forwarded message -----

From: <administrator@hrea.ca>
Date: Wed, Sep 28, 2022 at 8:53 AM
Subject: HREB - Approval of Ethics Renewal 587782
To: Hudson Jo-Anna(Principal Investigator) <j.hudson@mun.ca>
Cc: Moorehead Paul(Supervisor) <paul.moorehead@easternhealth.ca>, <administrator@hrea.ca>

Researcher Portal File #: 20190053

Dear Dr. Jo-Anna Hudson:

This e-mail serves as notification that your ethics renewal for study HREB # 2018.069 – Characterizing extracellular vesicles in acute lymphoblastic leukemia – has been **approved**. Please log in to the Researcher Portal to view the approved event.

Ethics approval for this project has been granted for a period of twelve months effective from **September 18, 2022** to **September 18, 2023**.

Please note, it is the responsibility of the Principal Investigator (PI) to ensure that the Ethics Renewal form is submitted prior to the renewal date each year. Though the Research Ethics Office makes every effort to remind the PI of this responsibility, the PI may not receive a reminder. The Ethics Renewal form can be found on the Researcher Portal as an "Event".

The ethics renewal **[will be reported]** to the Health Research Ethics Board at their meeting dated **October 6, 2022**.

Thank you,

Research Ethics Office

(e) info@hrea.ca

(t) 709-777-6974

(f) 709-777-8776

(w) www.hrea.ca

Office Hours: 8:30 a.m. – 4:30 p.m. (NL TIME) Monday-Friday

This email is intended as a private communication for the sole use of the primary addressee and those individuals copied in the original message. If you are not an intended recipient of this message you are hereby notified that copying, forwarding or other dissemination or distribution of this communication by any means is prohibited. If you believe that you have received this message in error please notify the

B2: *Ethics approval for effect of storage conditions on EV isolation study*

From: administrator@hrea.ca
Subject: HREB - Approval of Ethics Renewal 20222350
Date: January 10, 2023 at 8:39 PM
To: Christian Sherri Lynn(Principal Investigator) sherri@mun.ca
Cc: Pugh Trevor(External Co-PI) trevor.pugh@utoronto.ca, administrator@hrea.ca



Researcher Portal File #: 20222350

Dear Dr. Sherri Lynn Christian:

This e-mail serves as notification that your ethics renewal for study HREB # 2022.005 – Assessing storage conditions for extracellular vesicle isolation from urine and plasma – has been **approved**. Please log in to the Researcher Portal to view the approved event.

Ethics approval for this project has been granted for a period of twelve months effective from **January 25, 2023 to January 25, 2024**.

Please note, it is the responsibility of the Principal Investigator (PI) to ensure that the Ethics Renewal form is submitted prior to the renewal date each year. Though the Research Ethics Office makes every effort to remind the PI of this responsibility, the PI may not receive a reminder. The Ethics Renewal form can be found on the Researcher Portal as an “Event”.

The ethics renewal **will be reported** to the Health Research Ethics Board at their meeting dated **January 24, 2023**

Thank you,

Research Ethics Office

(e) info@hrea.ca

(t) 709-777-6974

(f) 709-777-8776

(w) www.hrea.ca

Office Hours: 8:30 a.m. – 4:30 p.m. (NL TIME) Monday-Friday

This email is intended as a private communication for the sole use of the primary addressee and those individuals copied in the original message. If you are not an intended recipient of this message you are hereby notified that copying, forwarding or other dissemination or distribution of this communication by any means is prohibited. If you believe that you have received this message in error please notify the original sender immediately.

Appendix C: Codes and scripts for bioinformatics

C1: Codes and scripts for chapter 3

```
# Differential expression analysis using z-scores with limma

setwd()

getwd()

#Script for normalizing and DE analysis of miRNA raw data #source: Limma User's guide (last
revision 12/06/2020)

library(oligo)

library(limma)

#To explore the list of GEO sample (GSM), GEO series (GSE) (list of GSM files that that make
up a single experiment)

gse <- getGEO("GSE56489", GSEMatrix = TRUE)

show(gse)

#Download raw data from GEO database into wd

filePaths = getGEOsuppFiles("GSE56489")

filePaths

#Because these have already been setup and optimised, don't run lines 10-16. start from line 19

#Set WD to the 'GSE56489_RAW2' folder
```

```

#Read in the raw files

#target.txt file describes the design of the experiment (denotes specific files for the data they
represent)

targets <- readTargets("GSE56489target7e.txt", row.names = "Name") #or name of targets file #
targets #to show target file components

febit <- list(E="VALUE")

other = c("ID_REF")

x <- read.maimages(targets, source="generic", columns=febit, annotation=other, green.only = T)

dim(x)

x #Note: X is an object of class "EListRaw"

#Replace NA with 0

summary(x)

x$E[is.na(x$E)] <- 0

x

boxplot(x$E)

#Quality control of unnormalized data

plotMD(x$E)

#Normalize between arrays

y <- normalizeBetweenArrays(x)

y #Note: y is an object of class "EList"

#Replace NA with 0

```



```
summary(y)
```

```
y$E[is.na(y$E)] <- 0
```

```
y
```

```
#plot boxplot
```

```
boxplot(log2(y$E), range=0, ylab="log2 intensity") #log2 transform only if that was not done by  
previous codes
```

```
#MD plot post norm
```

```
plotMD(y$E)
```

```
#MDS plot to show batch effect
```

```
#Multidimensional scaling (MDS) is a means of visualizing the level of similarity of individual  
cases of a dataset. MDS is used to translate "information about the pairwise 'distances' among a  
set of n objects or individuals" into a configuration of n points mapped into an abstract Cartesian  
space.
```

```
plotMDS(y$E)
```

```
y #Note: y is an object of class "EList"
```

```
#Differential gene expression analysis
```

```
#stipulate model matrix
```



```

#Open csv file in excel, delete x and additional column 1 and rows 1 and 2, save as csv and read
in
t.Zscores564 <- read.csv("t.Zscores564.2.csv", header = F)
t.Zscores564

#Calculate Z scores and write into csv
Z.scores.564 <- scale(t.Zscores564)
write.csv(Z.scores.564, file = "forzscores.1.csv")

#Add columns and rows and read in csv in the same order as in file from line 82
z.scores.564 <- read.csv("forzscores.2.csv")
z.scores.564
Z.scores.564 <- t(z.scores.564)
write.csv(Z.scores.564, file = "zscoresforheatmaps56489.csv")

# Venn diagram
library("gdata")
geneLists <- read.xls("vennsample1.xlsx", sheet=1, stringsAsFactors=FALSE, header=TRUE)
head(geneLists)
tail(geneLists)
geneLS <- lapply(as.list(geneLists), function(x) x[x != ""])
removeEMPTYstrings <- function(x) {

```

```

newVectorWOstrings <- x[x != ""]
return(newVectorWOstrings)

geneLS2 <- lapply(as.list(geneLists), removeEMPTYstrings)

# You can print the last 6 entries of each vector stored in your list, as follows:

lapply(geneLS, tail)

lapply(geneLS2, tail) # Both methods return the same results

# We can rename our list vectors

names(geneLS) <- c("ConditionA", "ConditionB")

# Now we can plot a Venn diagram with the VennDiagram R package, as follows:

require("VennDiagram")

VENN.LIST <- geneLS

venn.plot <- venn.diagram(VENN.LIST , NULL, fill=c("darkmagenta", "darkblue"),
alpha=c(0.5,0.5), cex = 2, cat.fontface=4, category.names=c("A", "B"), main="Random Gene
Lists")

# To plot the venn diagram we will use the grid.draw() function to plot the venn diagram

grid.draw(venn.plot)

# To get the list of gene present in each Venn compartment we can use the gplots package

```

```

require("gplots")

a <- venn(VENN.LIST, show.plot=FALSE)

# You can inspect the contents of this object with the str() function
str(a)

# By inspecting the structure of the a object created,
# you notice two attributes: 1) dimnames 2) intersections
# We can store the intersections in a new object named inters
inters <- attr(a,"intersections")

# We can summarize the contents of each venn compartment, as follows:
# in 1) ConditionA only, 2) ConditionB only, 3) ConditionA & ConditionB
lapply(inters, head)

#Volcano plots
res <- read.table("results.txt", header=TRUE)
head(res)

# Make a basic volcano plot
with(res, plot(log2FoldChange, -log10(pvalue), pch=20, main="Volcano plot", xlim=c(-2.5,2)))

# Add colored points: red if padj<0.05, orange if log2FC>1, green if both
with(subset(res, padj<.05 ), points(log2FoldChange, -log10(pvalue), pch=20, col="red"))

```

```
with(subset(res, abs(log2FoldChange)>1), points(log2FoldChange, -log10(pvalue), pch=20,
col="orange"))

with(subset(res, padj<.05 & abs(log2FoldChange)>1), points(log2FoldChange, -log10(pvalue),
pch=20, col="green"))

# Label points with the textxy function from the calibrate plot

library(calibrate)

with(subset(res, padj<.05 & abs(log2FoldChange)>1), textxy(log2FoldChange, -log10(pvalue),
labs=Gene, cex=.8))

res <- read.table("313.3_50.csv")

head(313)
```

C2: codes and scripts for chapter 4

```
library(limma)
```

```
library(DESeq2)
```

```
library(Rsubread)
```

```
library(readxl)
```

```
library(pheatmap)
```

```
library(ggplot2)
```

Set the working directory to 'plasma DESeq2 Nov 28'

```
setwd("~/Documents/plasma DESeq2 Nov 28")
```

```
#Create new matrix
```

```
plas_04 <- list.files(pattern = ".bam")
```

```
plas_04
```

```
plas_04_counts <- featureCounts(files = plas_04, annot.ext = "GRCh38_latest_genomic.gff",
```

```
isGTFAnnotationFile = TRUE, GTF.featureType = "gene", GTF.attrType = "gene")
```

```
write.table(x=data.frame(plas_04_counts$annotation[,c("GeneID")], plas_04_counts$counts,
```

```
stringsAsFactors = FALSE), file = "plas_04Jan.txt", quote = FALSE, sep = "\t", row.names =
```

```
TRUE)
```

Open txt file in excel format

Delete column 1

Add cell in column 1 row 1 and type Gene ID in new cell

“Control H” and replace “plasma_sorted.bam”, “_sorted.bam” with nothing.

Highlight row 1, “control H” and replace “X” with nothing

‘Control H’ and replace “HD” with “NCD”

‘Control H’ and add “-E” to “1442”, “1874”, “3701”, “4395” and “BALL10”

‘Control H’ and add “-p” to “3399”, “4353”, “4894”

Save as plas_04Feb.1.txt

```
plas_04matrix <- read.table("plas_04Feb.1.txt", sep = "\t", header = TRUE)
```

```
head(plas_04matrix)
```

```
n_plas04genes <- apply(plas_04matrix[,2:ncol(plas_04matrix)], 2, function(c)sum(c!=0))
```

```
write.table(n_plas04genes, file = "nplas04genes.txt", quote = F, row.names = TRUE, sep="\t")
```

```
filt_plas04genes <- apply(plas_04matrix[,2:ncol(plas_04matrix)], 1, function(x)length(x[x > 2])  
>=2)
```

```
plas04_filt_genes <- plas_04matrix[filt_plas04genes,]
```

```
head(plas04_filt_genes[,1:5])
```

```
write.table(plas04_filt_genes, file="plas04_filt_genes2.txt", quote = F, row.names = TRUE, sep  
= "\t")
```

```
design <- read_excel("expdes.xlsx")
```

```
design <- na.omit(design)
```

```
colData <- design
```

```
countData <- plas04_filt_genes
```

```
library("dplyr")
```



```

dds <- DESeqDataSetFromMatrix(countData, colData, design = ~Condition, tidy = TRUE)

dds

resultsNames(dds)

character(0)

norm <- rlog(dds,blind=FALSE)

head(assay(norm), 3)

norm_matrix <- assay(norm)

norm_df <- data.frame(Gene=rownames(norm_matrix), norm_matrix)

write.table(norm_df, "allplas_Jan13DESeq2 rlog.csv", quote = F, row.names = FALSE, sep =
"\t")

dds2 <- estimateSizeFactors(dds)

dds2

df2 <- bind_rows(as_data_frame(assay(norm)[, 1:2]) %>% mutate(transformation = "rlog"))
df2 <- bind_rows(as_tibble(assay(norm)[, 1:2]) %>% mutate(transformation = "rlog"))

colnames(df2)[1:2] <- c("x", "y")

lvls <- c("rlog")

df2$transformation <- factor(df2$transformation, levels=lvls)

vsd <- vst(dds, blind = FALSE)

head(assay(vsd), 3)

df3 <- bind_rows(as_tibble(assay(norm)[, 1:2]) %>% mutate(transformation = "rlog"),
as_data_frame(assay(vsd)[, 1:2]) %>% mutate(transformation = "vst"))

colnames(df3)[1:2] <- c("x", "y")

lvls <- c("vst", "rlog")

```

```

df3$transformation <- factor(df3$transformation, levels=lvls)

ggplot(df3, aes(x = x, y = y)) + geom_hex(bins = 80) + coord_fixed() + facet_grid( . ~
transformation)

sampleDists <- dist(t(assay(vsd)))

sampleDists

library("pheatmap")

library("RColorBrewer")

sampleDistMatrix <- as.matrix( sampleDists )

rownames(sampleDistMatrix) <- paste( vsd$Sample, vsd$Condition, sep = " - " )

colnames(sampleDistMatrix) <- NULL

colors <- colorRampPalette( rev(brewer.pal(9, "Blues")) )(255)

pheatmap(sampleDistMatrix,clustering_distance_rows = sampleDists,clustering_distance_cols =
sampleDists,col = colors)

#PCA

plotPCA(vsd, intgroup = "Condition")

plotPCA(vsd, intgroup = c("sample", "condition"))

library(tidyverse)

deg <- DESeq(dds)

res <- results(deg)

head(results(deg, tidy = TRUE))

```

```

summary(res)

res <- res[order(res$padj),]

head(res)

log2 fold change (MLE): Condition Disease vs control

Wald test p-value: Condition Disease vs control

DataFrame with 6 rows and 6 columns

write.table(res, file="plas13Jan_DPRNA.csv",quote=F, sep="\t", row.names = TRUE)

#To remove NA from padj

deg2 <- DESeq(deg, minReplicatesForReplace=Inf)

res3<- results(deg, cooksCutoff=FALSE, independentFiltering=FALSE)

write.table(res3,file="plas13_JanDPRNA.csv",quote=F, sep="\t", row.names = FALSE)

res4 <- results(deg,tidy=TRUE)

merged_res4 <- merge(norm_df,res4,by.x="Gene",by.y="row")

#MA plot

library("apeglm")

plotMA(res)

par(mfrow=c(2,3))

library(GOfuncR)

Input_Genes<-read.delim2("plas_04Feb.1.txt",header = T)

Go_Enrich_Out<- go_enrich(Input_Genes)

```

```

library(clusterProfiler)

search_kegg_organism('hsa', by='kegg_code')

hsa <- search_kegg_organism('Homo sapiens', by='scientific_name')

dim(hsa)

data(geneList, package="DOSE")

gene <- names(geneList)[abs(geneList) > 2]

kk <- enrichKEGG(gene = gene, organism = 'hsa', pvalueCutoff = 0.05)

head(kk)

kk2 <- gseKEGG(geneList = geneList, organism = 'hsa', minGSSize = 120, pvalueCutoff =
0.05, verbose = FALSE)

head(kk2)

mkk <- enrichMKEGG(gene = gene, organism = 'hsa', pvalueCutoff = 1, qvalueCutoff = 1)

head(mkk)

mkk2 <- gseMKEGG(geneList = geneList, organism = 'hsa', pvalueCutoff = 1)

head(mkk2)

get_wp_organisms()

data(geneList, package="DOSE")

gene <- names(geneList)[abs(geneList) > 2]

enrichWP(gene, organism = "Homo sapiens")

gseWP(geneList, organism = "Homo sapiens")

#Volcano plot

```

```

par(mfrow=c(1,1))
> with(res, plot(log2FoldChange, -log10(pvalue), pch=20, main="Volcano plot", xlim=c(-5,5)))
> with(subset(res, padj<.05 ), points(log2FoldChange, -log10(pvalue), pch=20, col="blue"))
> with(subset(res, padj<.05 & abs(log2FoldChange)>2), points(log2FoldChange, -log10(pvalue),
pch=20, col="red"))

```

#Z scores

Creating Z scores from rlog data

```
library(edgeR)
```

```
library(ggplot2)
```

```
library(limma)
```

Create 2 versions of the same excel spreadsheet

V 1= normal, with column and row names

V 2 = for z scores calculation, using Scale() function in R

Use the DERNA logs - plasNov28.2.DERNA

V 1= rlogplasNov28.DERNA.v1 (excel and txt)

V2 = rlogplasNov28.DERNA.v2 (txt, but work on it in excel)

Use V2

```
Plasrlog <- read.table("rlogplasNov28.DERNA.v2.txt", sep = "\t", header = FALSE)
```

```
clus <- t(scale(t(Plasrlog)))
```

Write out in excel

```
write.table(clus, file = "rlogplasNov28forzscores.txt", quote = F, row.names = FALSE, sep="\t")
```

Open the txt file V1, copy the column and row names

Open the txt file V2, insert a new column 1 and row 1 and paste the names copied from V1.

Save this new final version as "rlogplasNov28forzscores.2.txt" and use in genesis

Create a new version of the above file, and modify it, putting NCDs to the left of B-ALL samples. Save as rlogplasNov28forzscoresV2.txt

GSEA

```
library(clusterProfiler)
```

```
library(tidygraph)
```

```
library(tweenr)
```

```
library(pathview)
```

```
library(enrichplot)
```

```
library(ggplot2)
```

```
organism = "org.Hs.eg.db"
```

```
library(organism, character.only = TRUE)
```

```
setwd("~/Documents/sorted bam files plasma ")
```

```
df = read.csv("Plas13_Jan.DPRNA-GSEA2.csv", header = TRUE)
```

```
original_gene_list <- df$log2FoldChange
```

```
names(original_gene_list) <- df$Gene.ID
```

```
gene_list <- na.omit(original_gene_list)
```

```
gene_list = sort(gene_list, decreasing = TRUE)
```

```
keytypes(org.Hs.eg.db)
```

```
gse <- gseGO(geneList = gene_list, ont = "ALL", keyType = "ALIAS", nPerm = 10000,  
minGSSize = 3, maxGSSize = 800, pvalueCutoff = 0.05, verbose = TRUE, OrgDb = organism,  
pAdjustMethod = "none")
```

```
dotplot(gse, showCategory=10, split=".sign") + facet_grid(~.sign)
```

C3: codes and scripts for chapter 5

```
library(DESeq2)
```

```
library(limma)
```

```
library(Rsubread)
```

```
#Create matrix
```

```
lym <- list.files(pattern = ".bam")
```

```
lym
```

```
lym_counts <- featureCounts(files = lym, annot.ext = "GRCh38_latest_genomic.gff",
```

```
isGTFAnnotationFile = TRUE, GTF.featureType = "gene", GTF.attrType = "gene")
```

```
write.table(x=data.frame(lym_counts$annotation[,c("GeneID")], lym_counts$counts,  
stringsAsFactors = FALSE), file = "lym Jul 1.txt", quote = FALSE, sep = "\t", row.names =  
TRUE)
```

Open txt file in excel format

“Control H” and replace “bt2_” with nothing, and “_sorted.bam” with nothing

Name of document modified = lym Jul 1. xlsx

```
library(readxl)
```



```

lym_DE <- read.table("lym Jul 1.2.txt", sep = "\t", header = TRUE)
head(lym_DE)

n_lymgenes <- apply(lym_DE[,2:ncol(lym_DE)], 2, function(c)sum(c!=0))
write.table(n_lymgenes, file = "lym ngenes.txt",quote = F, row.names = TRUE, sep="\t")
lymfilt_genes <- apply(lym_DE[,2:ncol(lym_DE)], 1, function(x)length(x[x > 2]) >=2)
lym_filt_genes <- lym_DE[lymfilt_genes,]
head(lym_filt_genes[,1:5])

write.table(lym_filt_genes, file="lym_filt_genes.txt", quote = F, row.names = TRUE, sep = "\t")
design <- read.table("experdesignlym.txt", sep = "\t", header = T)

countData <- lym_filt_genes
design <- read_excel("experdesignlym.xlsx")

design <- na.omit(design)
colData <- design

dds <- DESeqDataSetFromMatrix(countData, colData, design = ~Treatment, tidy = TRUE)

norm <- rlog(dds,blind=FALSE)

norm_matrix <- assay(norm)
norm_df <- data.frame(Gene=rownames(norm_matrix), norm_matrix)

```

```

write.table(norm_df, "lym Jul1 DESeq2 rlog.csv", quote = F, row.names = FALSE, sep = "\t")

deg <- DESeq(dds)

res <- results(deg, tidy=TRUE)

merged_res <- merge(norm_df, res, by.x="Gene", by.y="row")

write.table(merged_res, file="lym Jul1.DERNA.csv", quote=F, sep="\t", row.names = FALSE)

pdf("lym Jul1 MAplot.pdf")

plotMA(deg, ylim=c(-5,5))

dev.off()

pdf("lym Jul1 pca.pdf")

plotPCA(norm, intgroup=c("Treatment"))

dev.off()

library("pheatmap")

ntd <- normTransform(deg)

select <- order(rowMeans(counts(deg, normalized=TRUE)), decreasing=TRUE)[1:25]

df <- as.data.frame(colData(deg)[,c("Treatment", "Sample")])

pheatmap(assay(norm)[select,], cluster_rows=FALSE, show_rownames=FALSE,
cluster_cols=TRUE, annotation_col=df)

dev.off()

#GSEA

library(clusterProfiler)

library(tidygraph)

```

```

library(tweenr)

library(pathview)

library(enrichplot)

library(ggplot2)

# we use ggplot2 to add x axis labels (ex: ridgeplot)

library(ggplot2)

# SET THE DESIRED ORGANISM HERE

organism = "org.Hs.eg.db"

BiocManager::install(organism, character.only = TRUE)

library(organism, character.only = TRUE)

setwd("/Users/mnl/Documents/sorted bam files lym/May 30th analysis")

# reading in data from deseq2

df = read.csv("gseacell+EVexclusive_jun29.csv", header=TRUE)

# we want the log2 fold change

original_gene_list <- df$log2FoldChange

# name the vector

names(original_gene_list) <- df$X

# omit any NA values

gene_list <- na.omit(original_gene_list)

# sort the list in decreasing order (required for clusterProfiler)

gene_list = sort(gene_list, decreasing = TRUE)

```

```
keytypes(org.Hs.eg.db)

gse <- gseGO(geneList = gene_list, ont = "ALL", keyType = "ALIAS", nPerm = 10000,
minGSSize = 3, maxGSSize = 800, pvalueCutoff = 1, verbose = TRUE, OrgDb = organism,
pAdjustMethod = "none")
```

```
require(DOSE)
```

```
dotplot(gse, showCategory=10, split=".sign") + facet_grid(.~.sign)
```

```
ridgeplot(gse) + labs(x = "enrichment distribution")
```

```
gseaplot(gse, by = "all", title = gse$Description[1], geneSetID = 1)
```

```
gseaplot(gse, by = "all", title = gse$Description[2], geneSetID = 2)
```

KEGG

```
ids<-bitr(names(original_gene_list), fromType = "ALIAS", toType = "ENTREZID",
OrgDb=organism)
```

```
dedup_ids = ids[!duplicated(ids[c("ALIAS")]),]
```

```
df2 = df[df$X %in% dedup_ids$ALIAS,]
```

```
df2$Y = dedup_ids$ENTREZID
```

```
kegg_gene_list <- df2$log2FoldChange
```

```
names(kegg_gene_list) <- df2$Y
```

```
kegg_gene_list<-na.omit(kegg_gene_list)
```

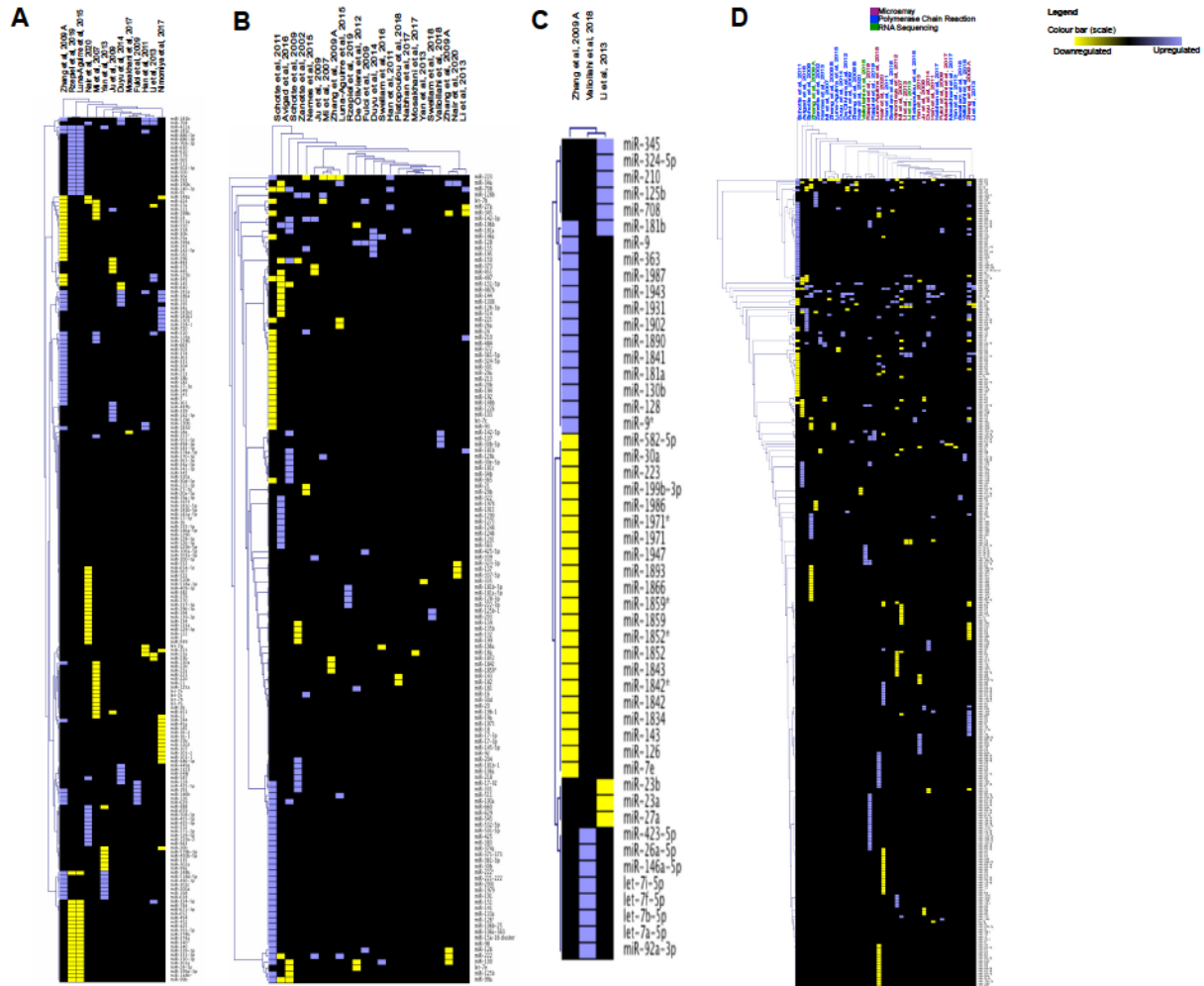
```
kegg_gene_list = sort(kegg_gene_list, decreasing = TRUE)
```

```
kegg_organism = "hsa"
```

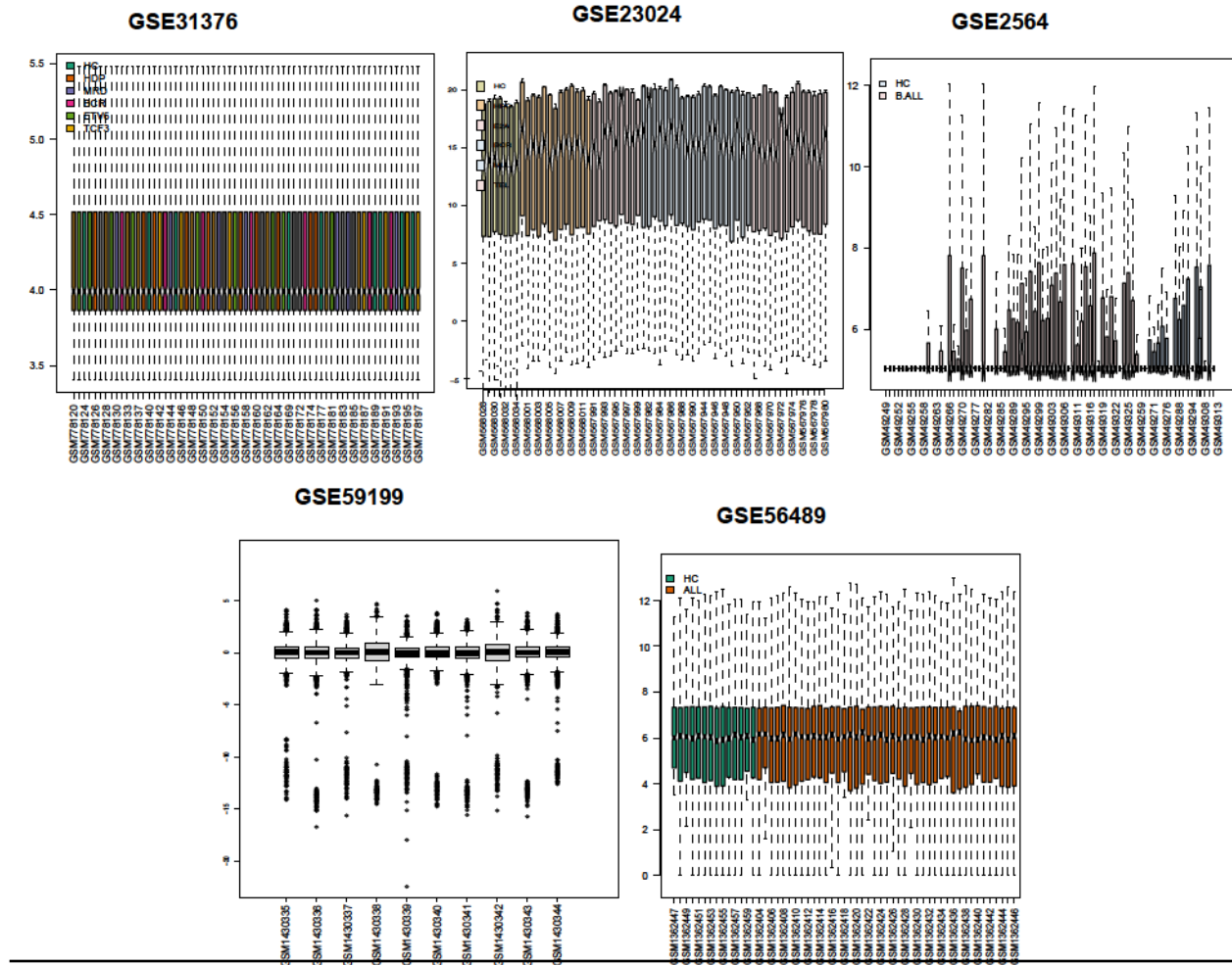
```
kk2 <- gseKEGG(geneList = kegg_gene_list, organism = kegg_organism, nPerm = 10000,  
minGSSize = 3, maxGSSize = 800, pvalueCutoff = 1, pAdjustMethod = "none", keyType =  
"ncbi-geneid")  
  
dotplot(kk2, showCategory = 10, title = "Enriched Pathways" , split=".sign") +  
facet_grid(.~.sign)  
  
ridgeplot(kk2) + labs(x = "enrichment distribution")  
  
gseaplot(kk2, by = "all", title = kk2$Description[1], geneSetID = 1)
```

Supplemental files and figures

Supplementary figures and files for miRNA signature of pediatric B-ALL (Chapter 3)



Supp. fig. 3.1. Average linkage hierarchical clusters for meta-analysis of scientific articles arranged by A. microarray, B. polymerase chain reaction, C. RNA Sequencing analyses, D. All three platforms merged. Horizontal bars were color-coded: yellow bars were downregulated DEmiRs, purple bars were upregulated DEmiRs while black bars were DEmiRs not shown in the study.



Supp. fig. 3.2. The raw data of healthy controls (HC) versus B-ALL were normalized and the normalized data were shown in boxplots.

Supp. Table 1. Table showing GEO ID, PMID, type of RNA analysis, miRNA analysis platform, B-ALL subtypes (when present), tissue source of samples, number of controls and patients and normalization techniques for reanalysis cohort datasets.

GEO Dataset	PMID	RNA analysis	Platform	B-ALL Subtypes	Age of patient	Tissue	Controls	Patients	Normalization
GSE31376	23888996	microarray	GPL8227 Agilent-019118 Human miRNA Microarray 2.0 G4470B (miRNA ID version)	Hyperdiploidy (12), T(9;22) (12), T(12;21)(14), T(1;19)(5), High MRD (9) - included. {T(1;1)(P34)(1), T(16;21)(1)- excluded	adult and pediatric	BM	13	65	quantile
GSE23024	21242186	RT-PCR	GPL10708 Applied Biosystems	10 MLL, 14 TEL- AML1, 10 BCR- ABL, 9 E2A-	Pediatric	BM	7	70	quantile

			TaqMan MicroRNA array (v 1.0, early Access) and GPL10709 Erasmus MC- Sophia Children's Hospital Custom TaqMan Array	PBX1, 13 hyper diploid, 14 unclassified B- ALL					
GSE2564	15944708	microarray	GPL1986 and GPL1987 Bead- based microRNA profiling	B-ALL	Pediatric	B-ALL	6	80	quantile

			platform versions 1 and 2						
GSE59199	32128558	microarray	GPL17107 Exiqon miRCURY LNA microRNA Array	B-ALL	adult and pediatric	normal blood	5	5	Print tip loess and scale
GSE56489	24955371	microarray	GPL14132 Febit Homo Sapiens miRbase 15.0	B-ALL	Pediatric	PB and BM	14	34	array and quantile

Supp. Table 2. Table showing list of scientific papers used for meta-analysis, with rows showing article, PMID, patient age classification, cytogenetic subtype (when known) and molecular subtypes (when known) and the RNA analyses platforms.

	Article	PMID	Adult vs. pediatric	PB vs. BM	Cytogenetic Subtypes	Molecular Subtypes	RNA Analysis platform
1	Mi et al, 2007	18056805		PB and BM		<p>one t(4;11)/<i>MLL-AF4</i> (<i>MLL-AFF1</i>) and two t(11;19)(q23;p13.3)/<i>MLL-ENL</i> (<i>MLL-MLLTI</i>) and seven ALL cell lines [i.e., two t(4;11) and five <i>MLL-ENL</i></p>	Microarray

2	Ju et al, 2009	19206004	pediatric	BM	CD19+ CD22+/- (Pre-B)	-	Microarray
3	Schotte et al, 2009	18923441	pediatric	PB and BM	CD19+ Pre- B (57)	MLL rearranged and B Other (19), hyperdiploidy (10), ETV6::RUNX1 (10), BCR::ABL (10), E2A::PBX1 (8)	
4	De Oliveira et al, 2012	22099053	pediatric	BM	Pro-B (10), CALLA/Pre- B (98),		
5	Duyu et al, 2014	24955371	pediatric	PB and BM			Microarray

6	Zhang et al, 2009 A	19915715	pediatric	BM			Microarray
7	Zhang et al, 2009 B	19724645	age not stated	BM			
8	Mosakhani et al, 2017	29068867	pediatric	BM		BCR::ABL1(1), MLL (2), del9p (6), ETV6::RUNX1 (20), Trisomy (3), Hyperdiploid (20), B other (20)	Microarray
9	Nabhan et al, 2017	28732737	pediatric	PB		Pre-B (23).B (5)	
10	Fulci et al, 2009	19760605	adult			BCR::ABL1 (17), TCF::PBX1 (4),	Microarray

						MLL (7), No mutations (15)	
11	Schotte et al, 2011	21242186	pediatric	PBMCs, 7 normal BM		MLL (10), ETV6::RUNX1 (14), BCR::ABL1 (10), TCF::PBX1 (9), hyperdiploid (13), B Other (14)	
12	Nemes et al, 2015	25388103	pediatric		pre-B (15) normal B cells = CD19+		Microarray
13	Yan et al, 2013	23888996	pediatric	PB		ETV6::RUNX1 (15), BCR::ABL1 (12), TCF::PBX1	

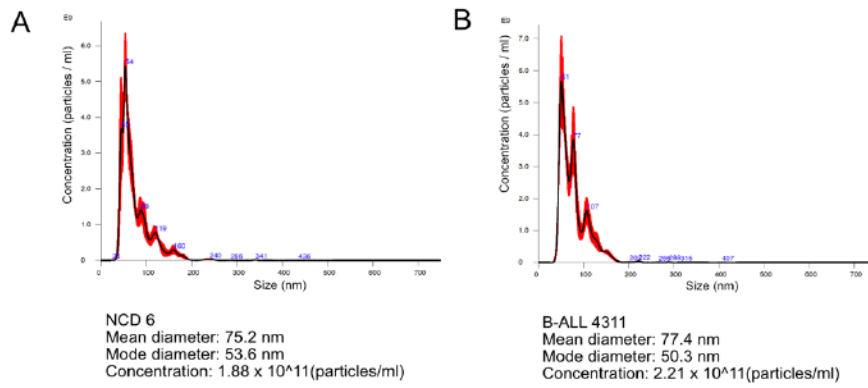
						(4), Hyperdiploid (12), B Other (20)	
14	Han et al, 2011	21926415	pediatric	BM	pre-B (11), pro-B, CALLA		
15	Avigad et al, 2016	26684414	pediatric	BM			
16	Piatopoulou et al, 2018	29556721	pediatric	BM	CD19+ CD20+, CD22+, CD24+, CD10+, CD34+, CD38+,	High hyperdiploid (yes = 22, no=102, unknown = 1), hypodiploidy (yes = 4, no = 120, unknown =1), ETV6::RUNX1 (No = 93, yes = 31,	Microarray

					CD58+, CD11a,	unknown = 1), BCR::ABL1 (no = 122, yes = 2, unknwon = 1)	
17	Swellem et al, 2016	26857279	pediatric		pre-B (32)		Microarray
18	Swellam et al, 2018	29460192	pediatric	PB	pre-B (25)	-	Microarray
19	Nair et al, 2020*	32531485	pediatric	PB and BM		Ph-negative (24)	
20	Rzepiel et al, 2019	31727091	pediatric	PB (at diagnosis, days 8, 15 and 33) and BM	pre-B (28 de novo, 5 at 1st relapse) (note; samples	de novo = hyperdiploidy (3), ETV6::RUNX1 (6), no known mutations (6)	

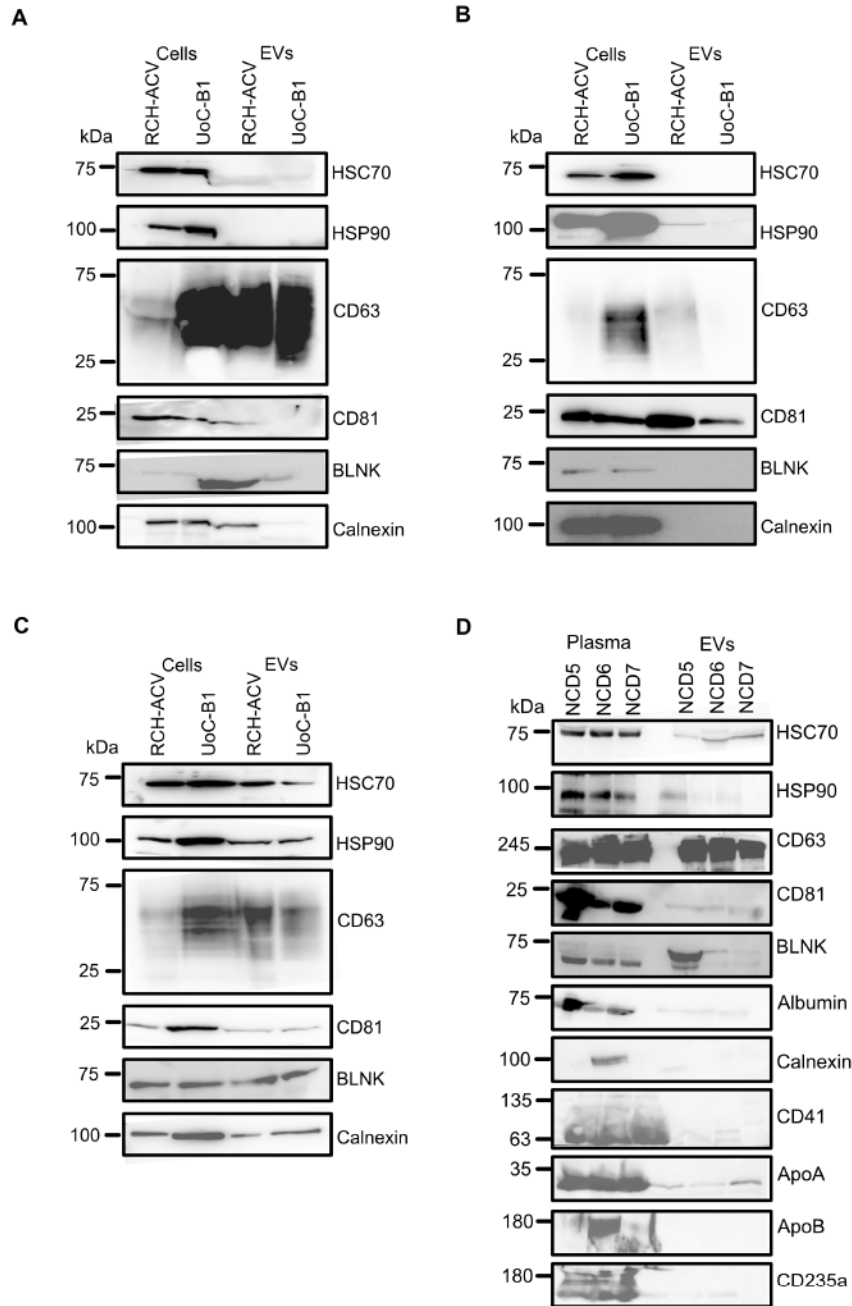
				(at diagnosis days 15 and 33)	collected at different time points)		
21	Zanette et al, 2002	17934639	pediatric	PB and BM	Pro-B , CALLA,	-	
22	Luna-Aguirre et al, 2015	26406572	adult (11) vs pediatric (19)	PB	Early Pre B (1), CALLA (13), Pre-B (11), B-ALL (14)	-	Microarray
23	Valiollahi et al, 2018	30712246	adult	BM	CD34+, CD38- = LSC,	BCR-ABL	Microarray

					CD34+, CD38+ = LP		
24	Li et al, 2013	23611221	pediatric	BM	CALLA (Pre-B) = CD10+, no surface or cytoplasmic Ig	BCR-ABL, MLL- AF4 and non- subtyped	
25	Ninomiya et al, 2012	22456238	pediatric	PB and BM		BCR-ABL 1+, BCR- ABL-	Microarray

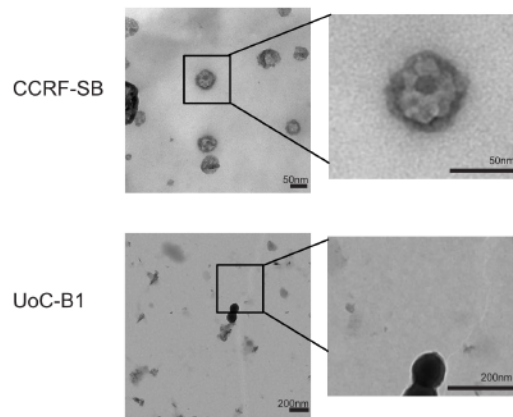
Supplementary figures for EV-based RNA signature of pediatric B-ALL (Chapter 4)



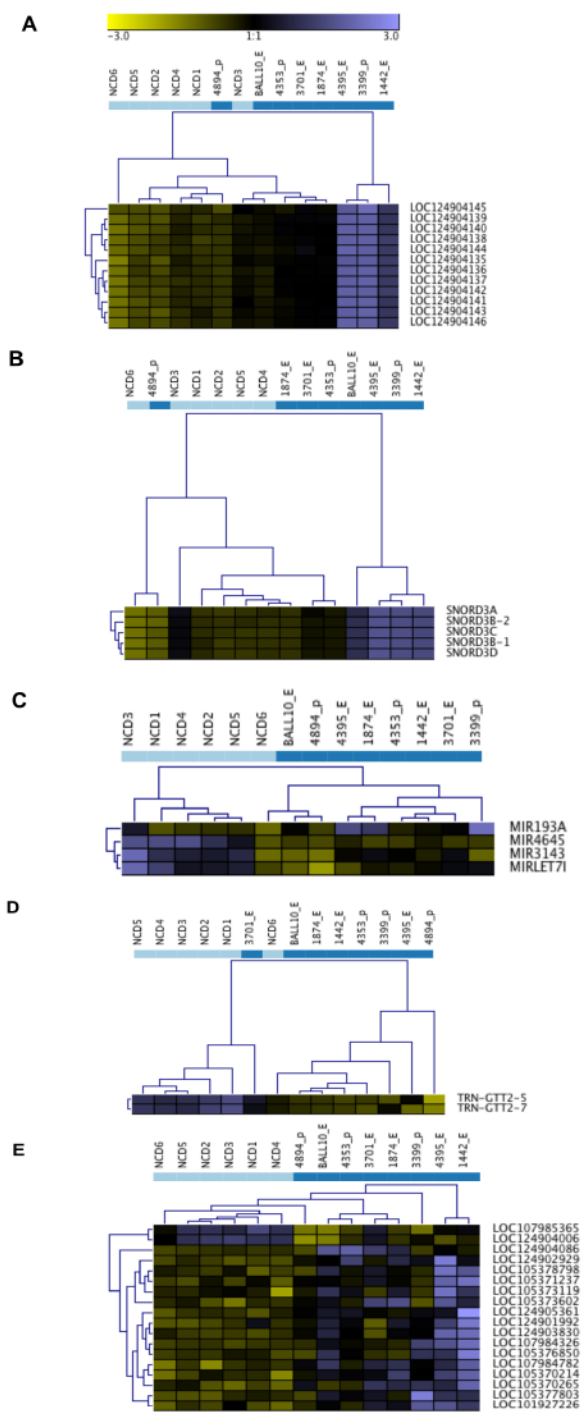
Supp. fig. 4.1. Sample NTA data for NCD and pediatric B-ALL samples, showing histogram – typical output of NS300 run.



Supp. fig 4.2. Western blot image comparing EV isolation techniques.

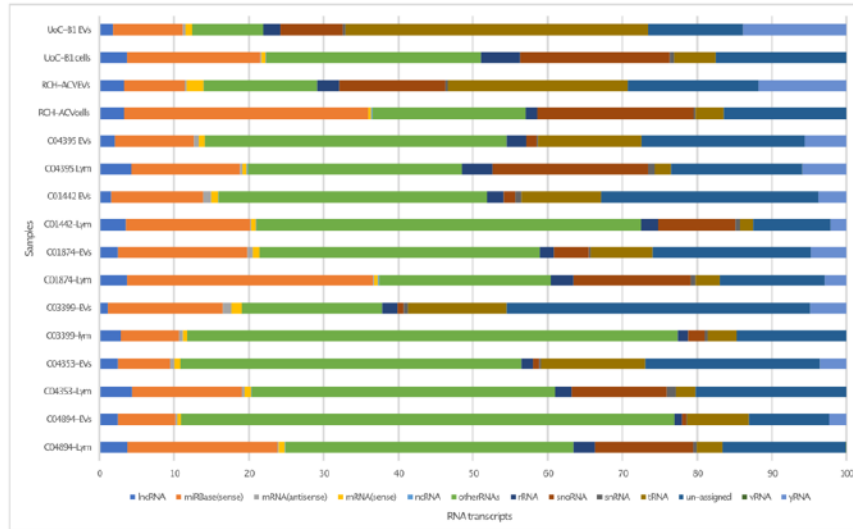


Supp. fig. 4.3. Transmission electron microscopy (TEM) images of pediatric B-ALL cell line EVs from CCRF-SB cell line (top) and UoC-B1 (bottom) isolated with Vn96.

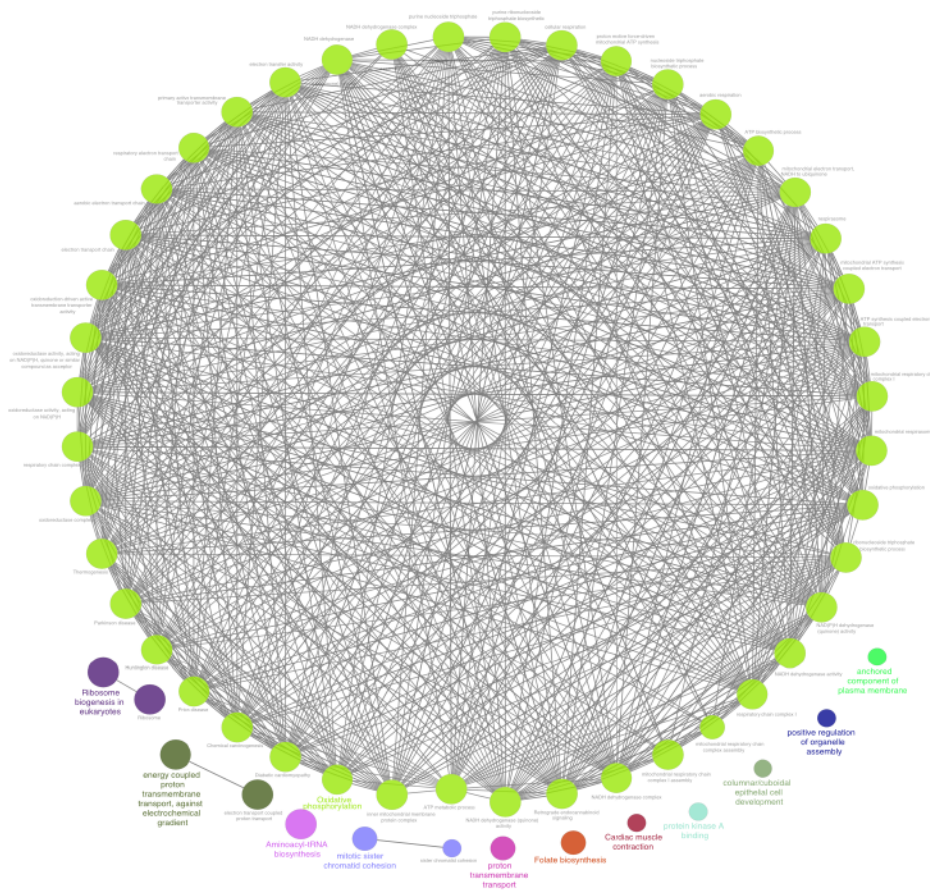


Supp. fig. 4.4. Unsupervised hierarchical cluster of small RNA EV clusters that are not clearly discriminatory.

Supplementary figures for RNA sorting mechanisms in pediatric B-ALL (chapter 5)



Supp. fig 5.1. RNA transcript types in cell and EV samples.



Supp. fig. 5.2. Gene ontology (GO) and Kyoto encyclopedia of gene and genomes (KEGG) pathways for top 200 non-selectively packaged RNAs, based on average raw counts.