Health Research Symposium 2024

Long noncoding RNAs in acute myeloid leukaemia
Roles for classification and prognostication

IP Ho-Wan Alvin
Department of Pathology
Queen Mary Hospital
26 November 2024

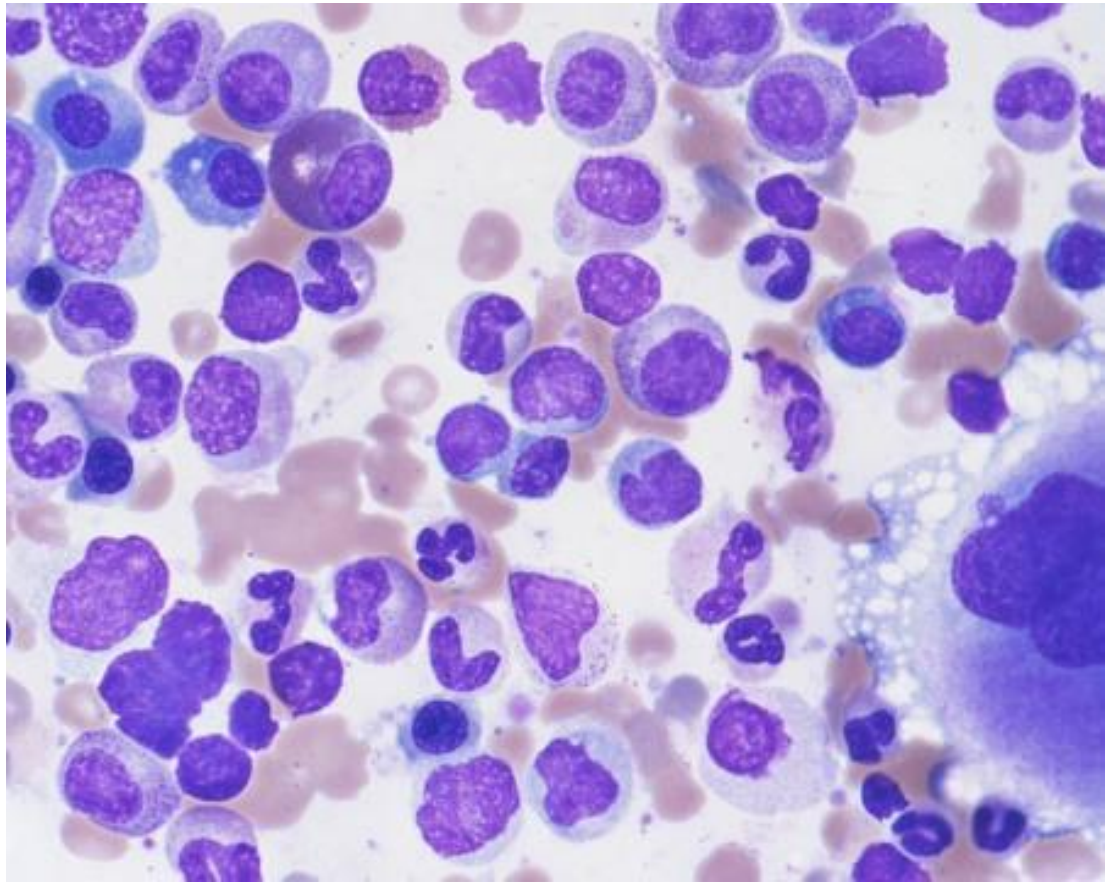
Outline

HMRF project (05160046)

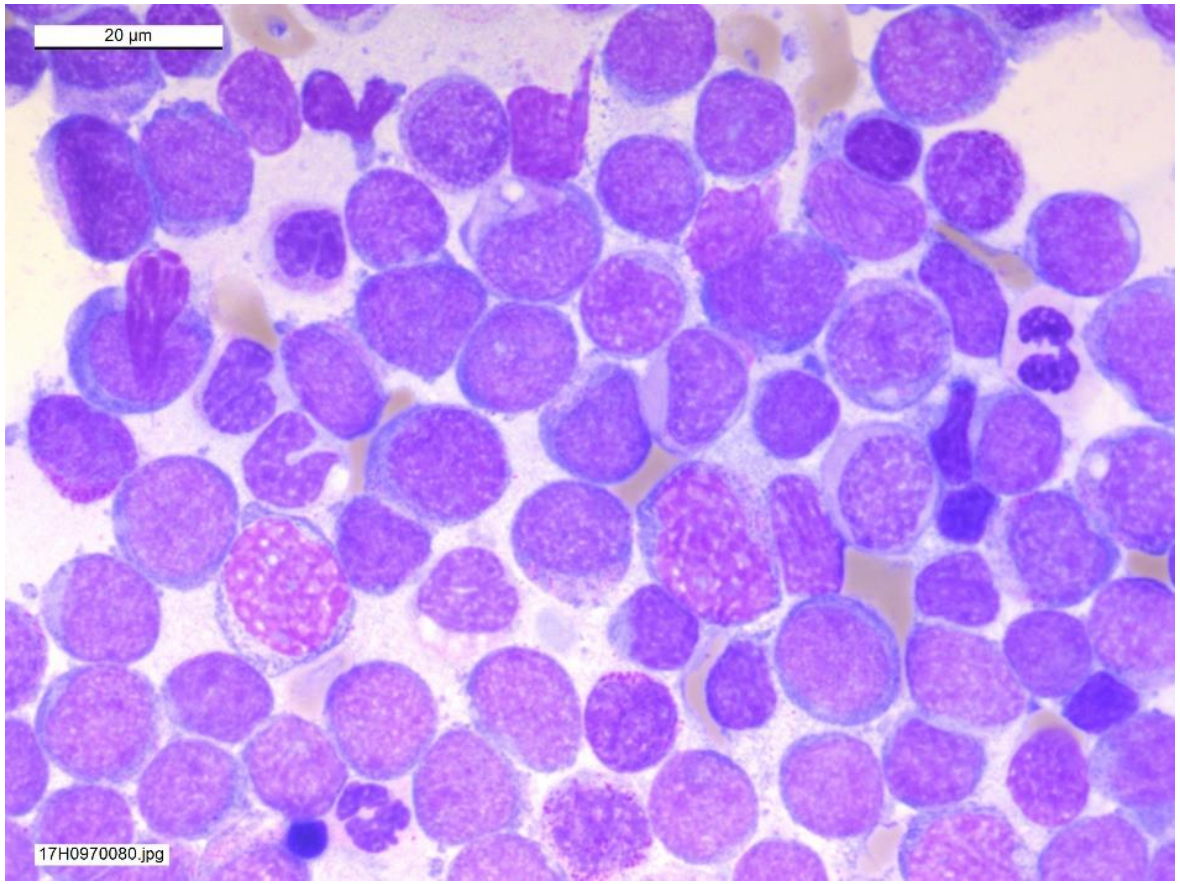
- Brief background on **acute myeloid leukaemia (AML)** and **long noncoding RNAs (lncRNAs)**
- **Prognostication** in AML
- Use of machine learning to devise **lncRNA prognostic score**
- **Validation and clinical application** of lncRNA prognostic score

Acute myeloid leukaemia (AML): Clonal disorder of myeloid stem/progenitor cells

Normal

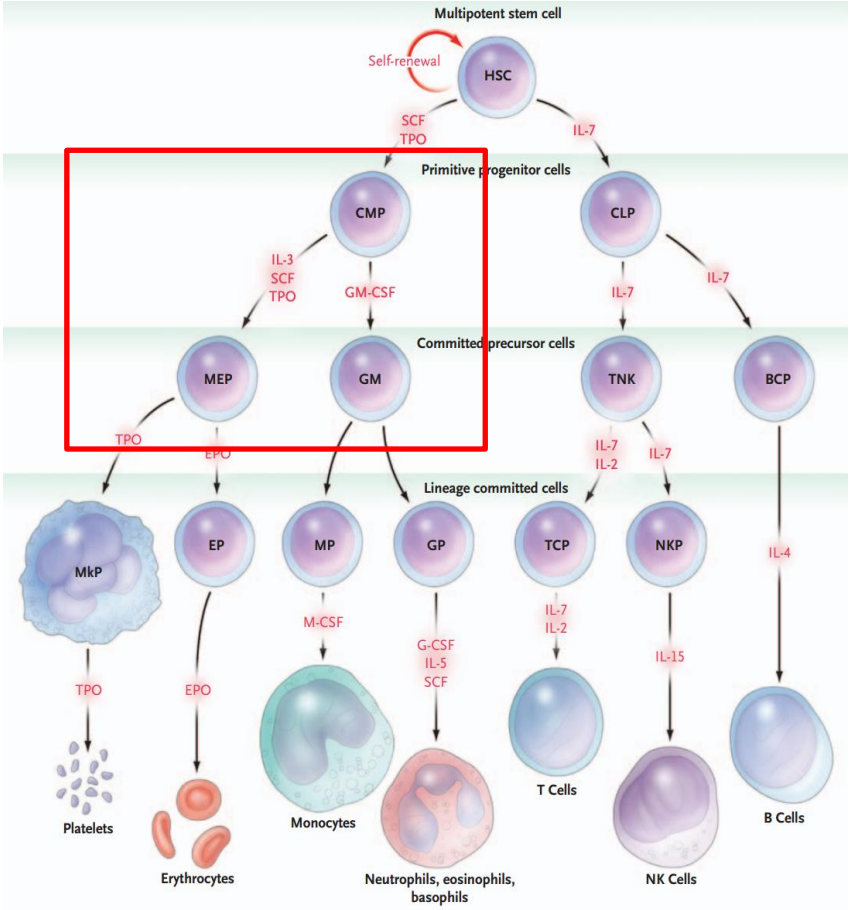


AML



Acute myeloid leukaemia
Brief introduction

- Clonal disorder of myeloid stem/progenitor cells
- ↑ proliferation
- ↓ differentiation
- **Expansion of immature/leukaemic cells** results in **severe reduction in normal blood cell production**, resulting in reduced normal blood cells in circulation
- Clinical problems: Anaemia, infections, bleeding



Kaushansky K. *N Engl J Med* 2006;354:2034-2045.

Acute myeloid leukaemia

Critical information for patient management

- **Classification**
  - Subtyping permits understanding of disease behaviour and better prediction of outcome
- **Prognostication**
  - Identify features to predict outcome
- **Treatment**
  - Select targeted therapies
- **Monitoring**
  - Identify target for post-treatment monitoring

Classification of AML

WHO 2022 Classification

Acute myeloid leukaemia with defining genetic abnormalities

- Acute promyelocytic leukaemia with *PML::RARA* fusion
- Acute myeloid leukaemia with *RUNX1::RUNX1T1* fusion
- Acute myeloid leukaemia with *CBFB::MYH11* fusion
- Acute myeloid leukaemia with *DEK::NUP214* fusion
- Acute myeloid leukaemia with *RBM15::MRTFA* fusion
- Acute myeloid leukaemia with *BCR::ABL1* fusion
- Acute myeloid leukaemia with *KMT2A* rearrangement
- Acute myeloid leukaemia with *MECOM* rearrangement
- Acute myeloid leukaemia with *NUP98* rearrangement

Fusions

- Acute myeloid leukaemia with *NPM1* mutation
- Acute myeloid leukaemia with *CEBPA* mutation
- Acute myeloid leukaemia, myelodysplasia-related

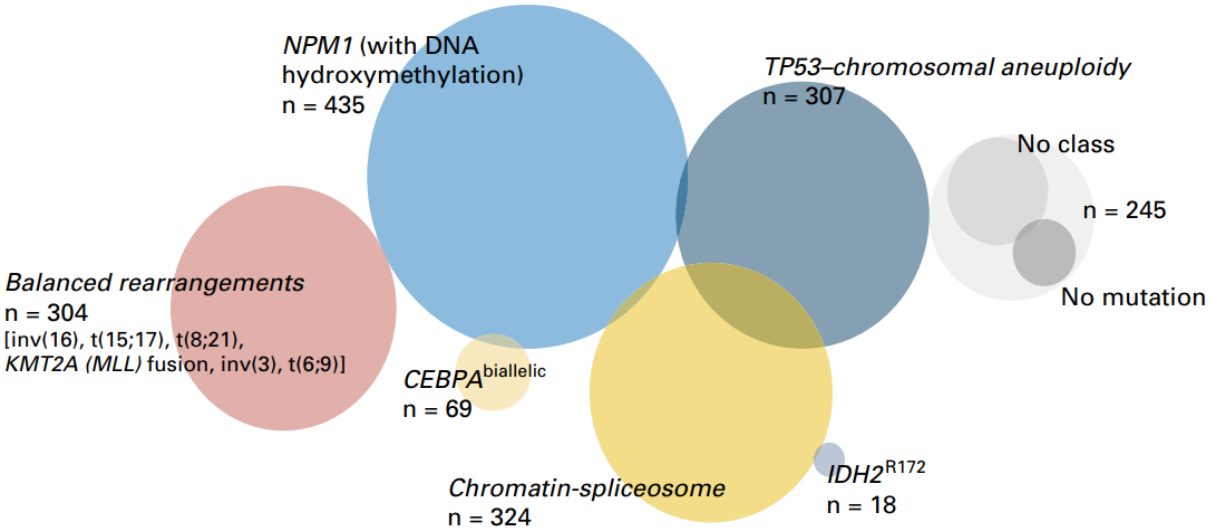
Mutations

- Acute myeloid leukaemia with other defined genetic alterations

Acute myeloid leukaemia, defined by differentiation

- Acute myeloid leukaemia with minimal differentiation
- Acute myeloid leukaemia without maturation
- Acute myeloid leukaemia with maturation
- Acute basophilic leukaemia
- Acute myelomonocytic leukaemia
- Acute monocytic leukaemia
- Acute erythroid leukaemia
- Acute megakaryoblastic leukaemia

Morphology



Khoury JD et al. *Leukemia* 2022;36:1703-1719.
Bullinger L et al. *J Clin Oncol* 2017;35:934-946.

Prediction of clinical outcome in AML

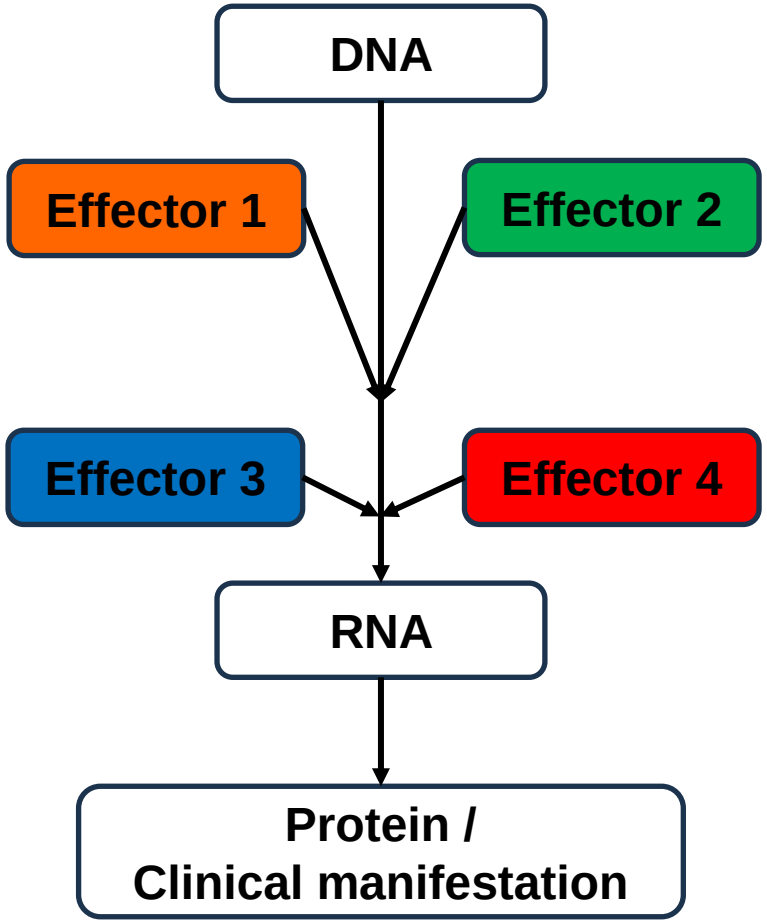
ELN risk classification

Risk category†	Genetic abnormality
Favorable	• t(8;21)(q22;q22.1)/*RUNX1*::*RUNX1T1*†,‡ • inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/*CBFB*::*MYH11*†,‡ • Mutated *NPM1*†,§ without *FLT3*-ITD • bZIP in-frame mutated *CEBPA*‖
Intermediate	• Mutated *NPM1*†,§ with *FLT3*-ITD • Wild-type *NPM1* with *FLT3*-ITD (without adverse-risk genetic lesions) • t(9;11)(p21.3;q23.3)/*MLLT3*::*KMT2A*†,¶ • Cytogenetic and/or molecular abnormalities not classified as favorable or adverse
Adverse	• t(6;9)(p23.3;q34.1)/*DEK*::*NUP214* • t(v;11q23.3)/*KMT2A*-rearranged# • t(9;22)(q34.1;q11.2)/*BCR*::*ABL1* • t(8;16)(p11.2;p13.3)/*KAT6A*::*CREBBP* • inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)/*GATA2*, *MECOM*(*EVI1*) • t(3q26.2;v)/*MECOM*(*EVI1*)-rearranged • −5 or del(5q); −7; −17/abn(17p) • Complex karyotype,** monosomal karyotype†† • Mutated *ASXL1*, *BCOR*, *EZH2*, *RUNX1*, *SF3B1*, *SRSF2*, *STAG2*, *U2AF1*, and/or *ZRSR2*‡‡ • Mutated *TP53*[a]

- Latest addition of chromatin/spliceosome gene mutations in the adverse risk group has enabled prognostication in 15-20% more AML patients
- Is this optimal in predicting patients' outcome?

Döhner H et al. *Blood* 2022;140:1345-1377.

Genome vs. Transcriptome (DNA vs. RNA)



- DNA: Only detects genomic variants (spelling mistakes)
- RNA: Final common pathway that captures more effectors (or “epigenomics”)

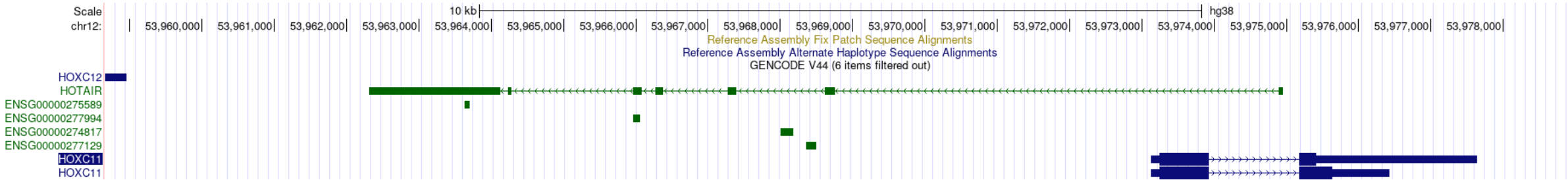
ATGCTGCATCGATCGATCGATCGATCGTACTACGATCGATCGATCGATCGATTGTGTGCA
TCGTACGACTGATCGATCGATCGTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT
AGCTAGCTAGCATCGATCGATCGATCGATCGATCGTACGATCGATCGATCGATCGTACGA
TCGATCGATCGATCGATCGTACGATCGTACGTACGTACGTACGTACGTACGATCATCGAT
CGATCGATCGATCGATCGATCATCGTAGCATCACAACACATCGATCGTCGTACATCATCA
TCATGCTAGCTAGCTAGCTAGCTAGCTAGCTGATCGATCGTACGTAGCATGCTACGTAGC
TAGCTAGCTAGCTCGTACGTAGCTACGATCGACATCATCGTACGTAGCTACGTACGTACG
TAGCTACGTACGTACGTACGTAGCACCACAATCGATCGTAGCTAGCTAGCTGTACGTAGC
TAGCTAGTCGTAGCTAGCTACGTAGCTACGTACATGCAGTCAGTCAGTCGATATGC

Holistic picture of the human pan-genome

https://blog.createandcraft.tv/layered-papercut-art-template/

Long noncoding RNAs

Very brief introduction

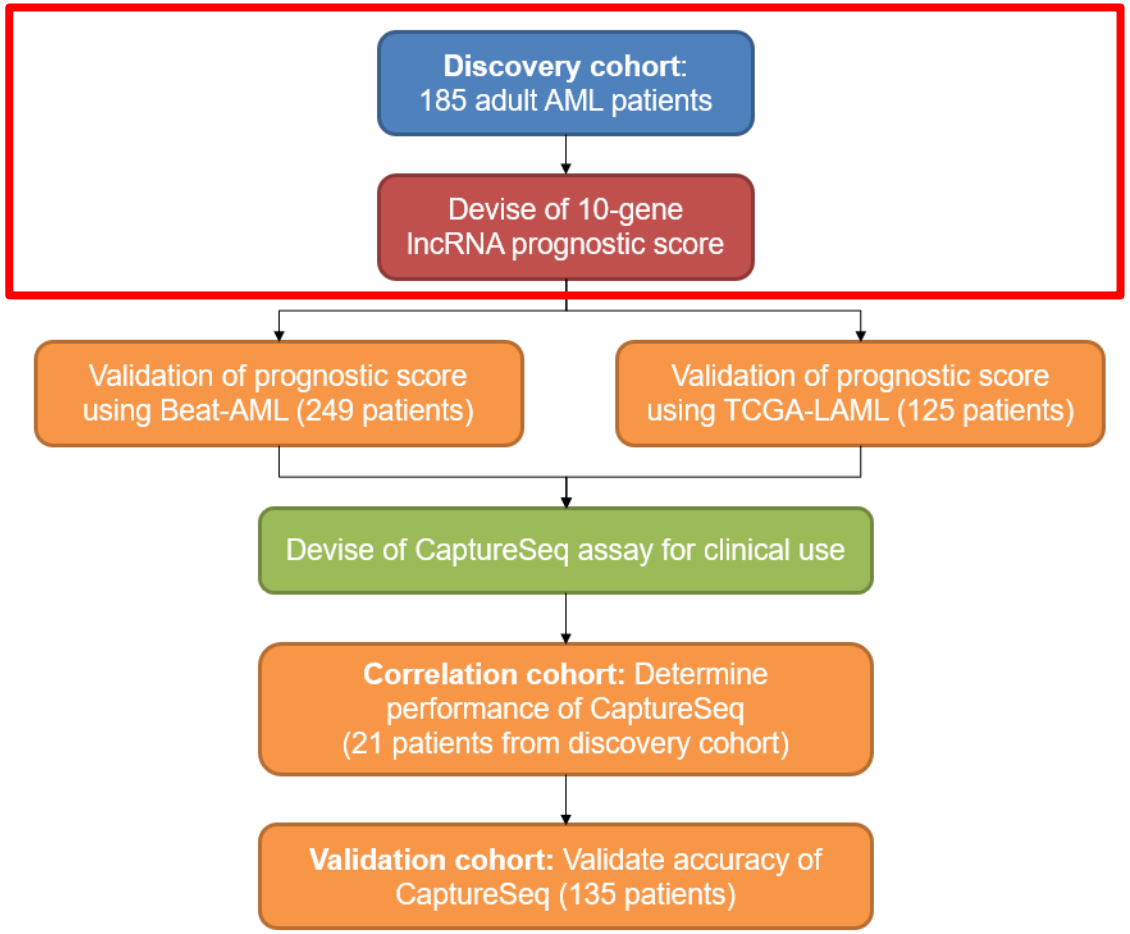


- Elements of the transcriptome
  - Messenger RNAs (protein-coding genes): 2% of human genome
  - Long noncoding RNAs (lncRNAs)
  - And many more RNA subtypes…

- lncRNAs: Transcripts longer than 200 nucleotides that do not appear to have a protein-coding sequence
- Over 100,000 lncRNAs recorded in human (cf. ~20,000 protein-coding genes)

Long noncoding RNAs in AML
Experimental design

- Project period: 2018-2022
- 185 adult patients with newly diagnosed AML for deep total transcriptome sequencing (between period of 2007 to 2018)
- Median follow-up: 417 days
- Detect only established lncRNAs to study clinical outcome



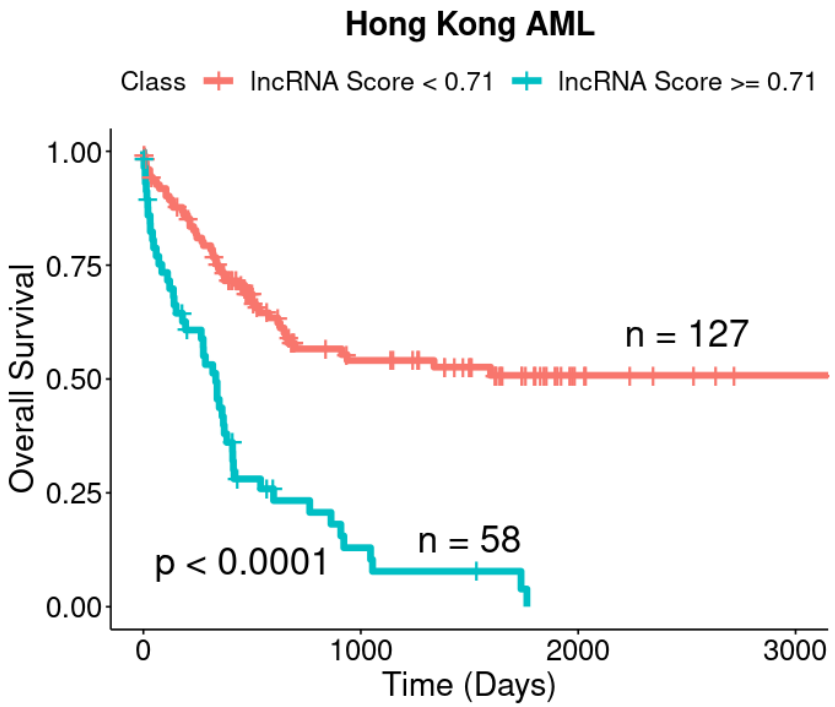
Long noncoding RNAs in AML
Classification

- Unsupervised clustering
- Classify established AML subtypes largely in accordance with their diagnostic categories
- “Blocks of colours” imply good classification of AML subtypes using lncRNAs

Long noncoding RNAs in AML

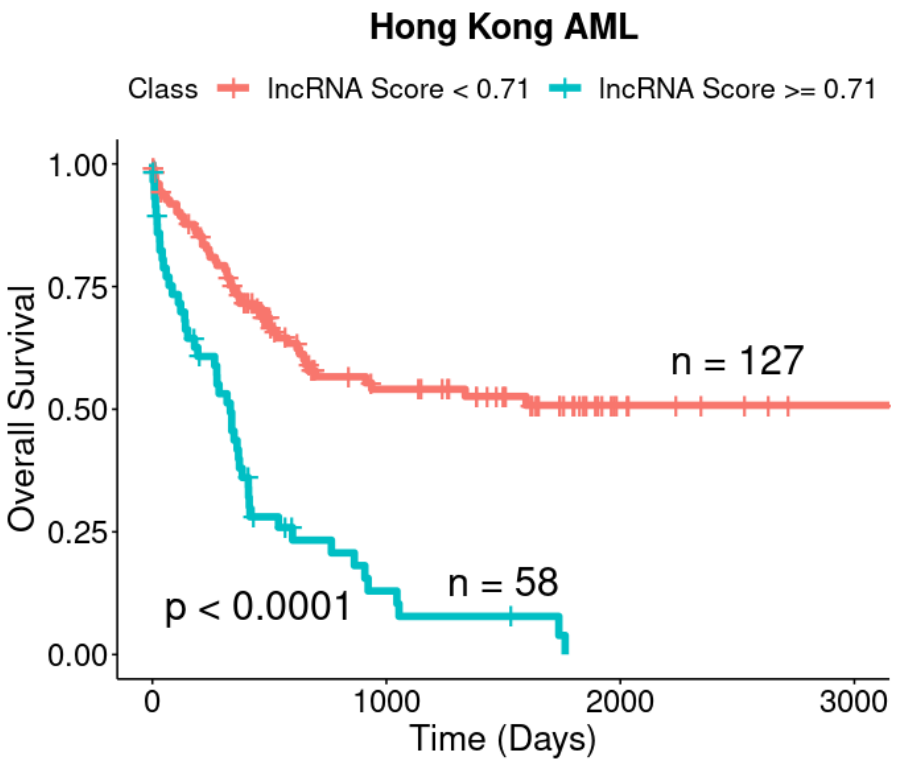
Identification of prognostic lncRNAs

- Machine learning: Lasso regression
- Identified 10 lncRNAs
- lncRNA prognostic score calculated for each patient:
  - Multiply 10 lncRNA expression level with their corresponding weighted coefficient from Lasso
  - Linearly combining their products



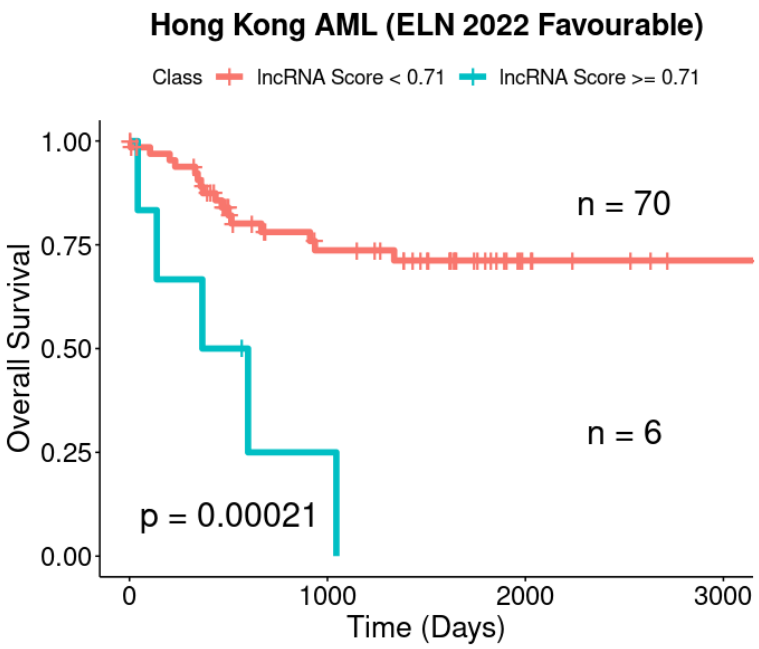
Long noncoding RNAs in AML

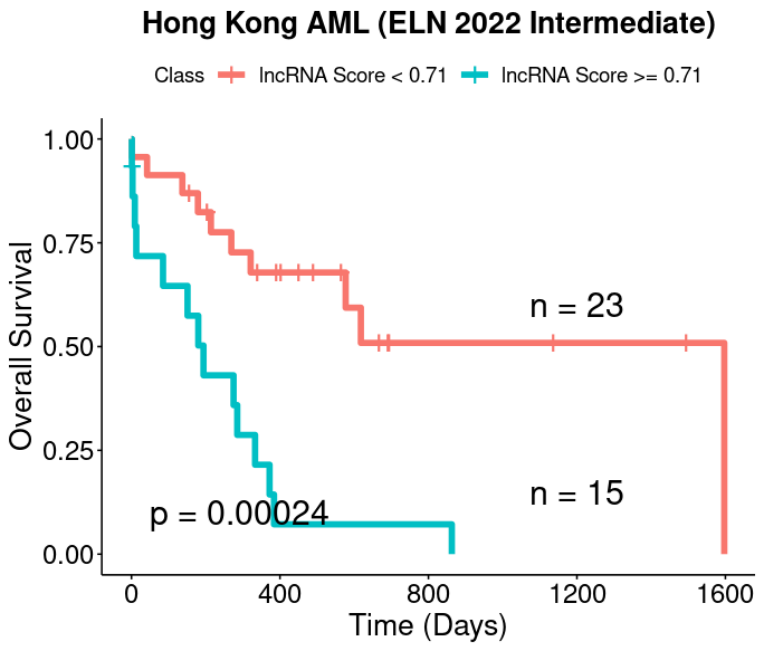
Multi-variable analysis of prognostic effects in HK cohort

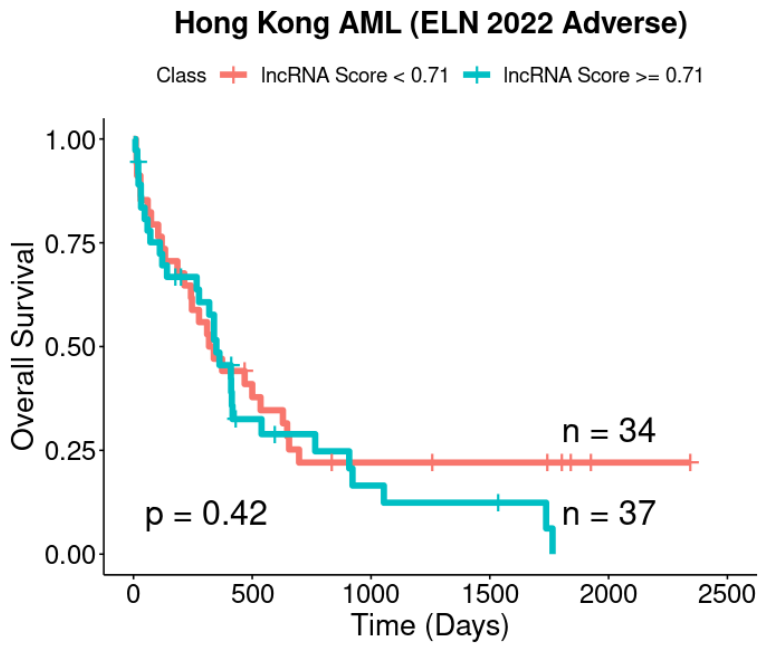


Prognostic value of lncRNAs in AML

Interactions with current prognostic system

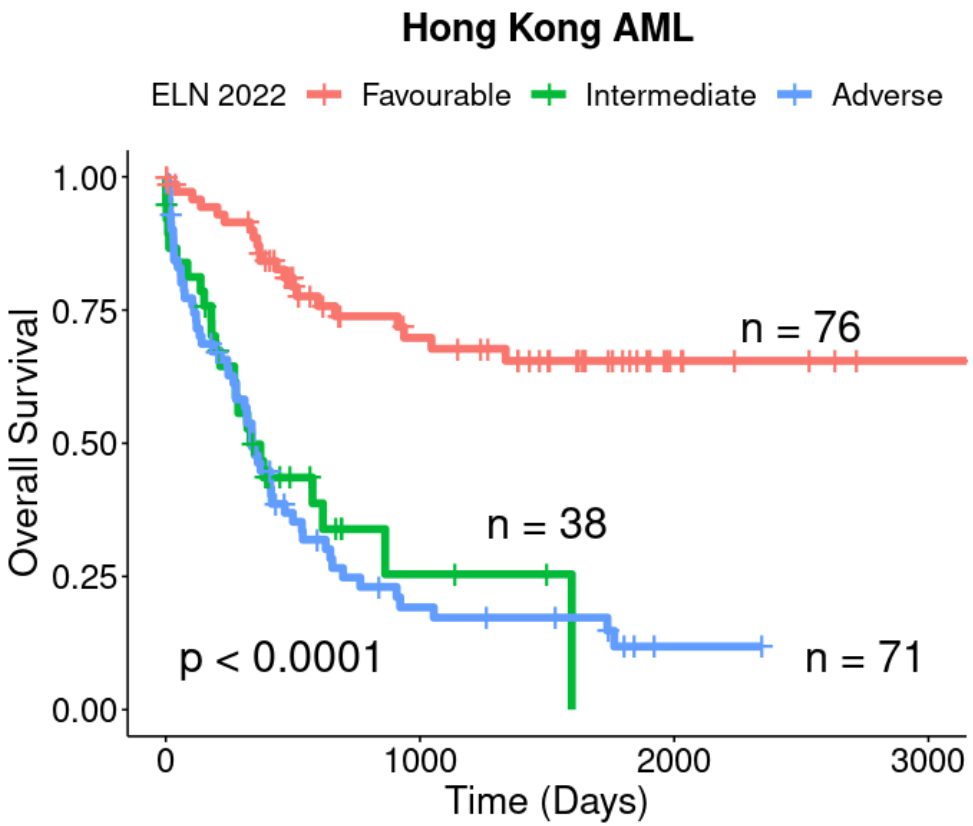


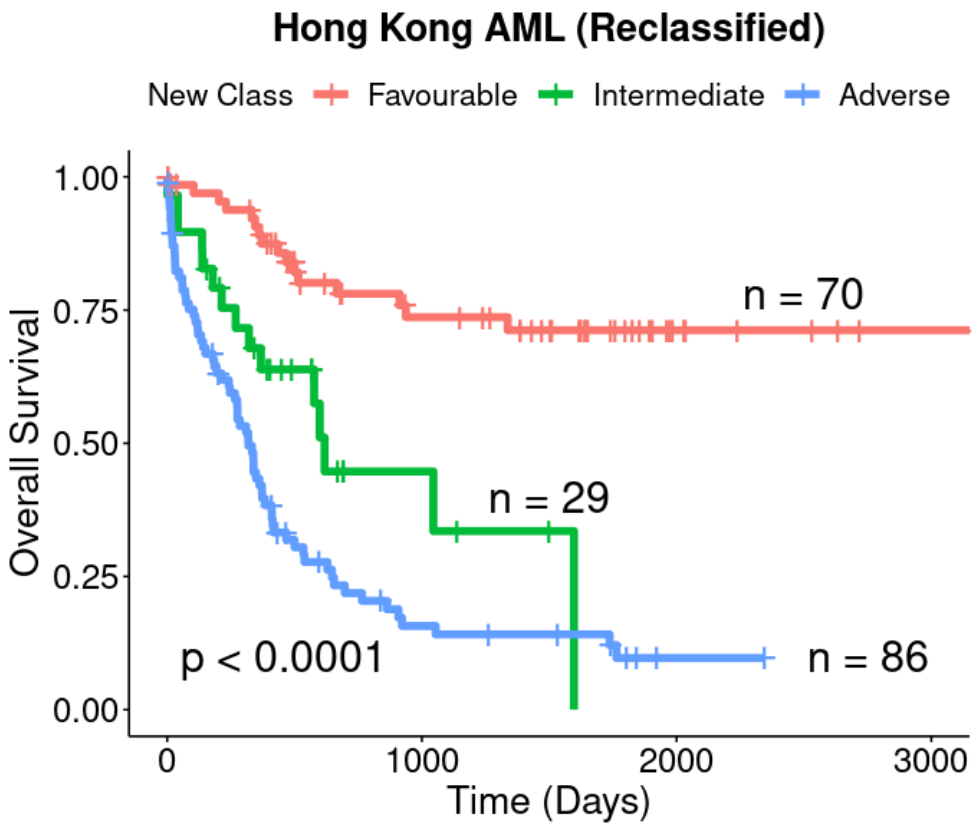




Prognostic value of lncRNAs in AML

Interactions with current prognostic system





Validation of 10-lncRNA score

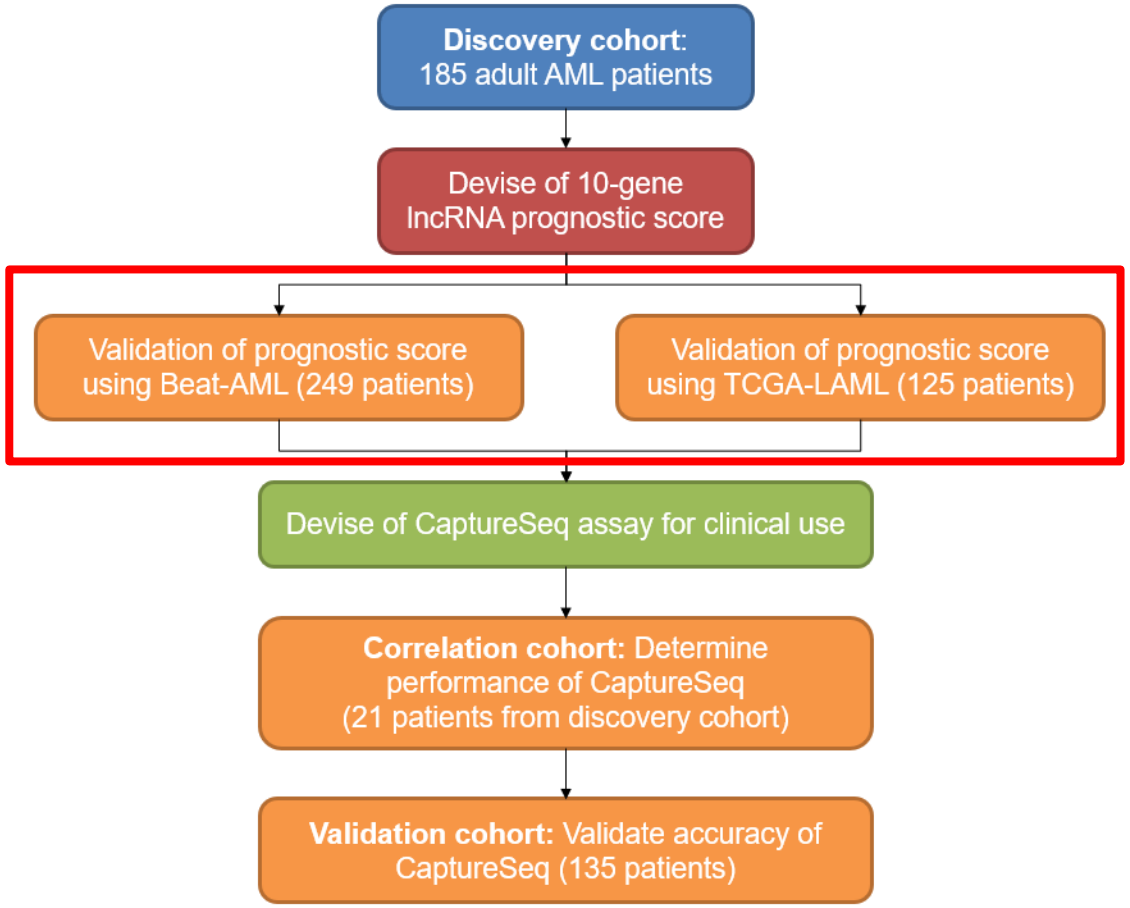
Role of external data sets

- Two well-established data sets with available
  - Transcriptome data
  - Survival data

- The Cancer Genome Atlas (TCGA)
- BeatAML

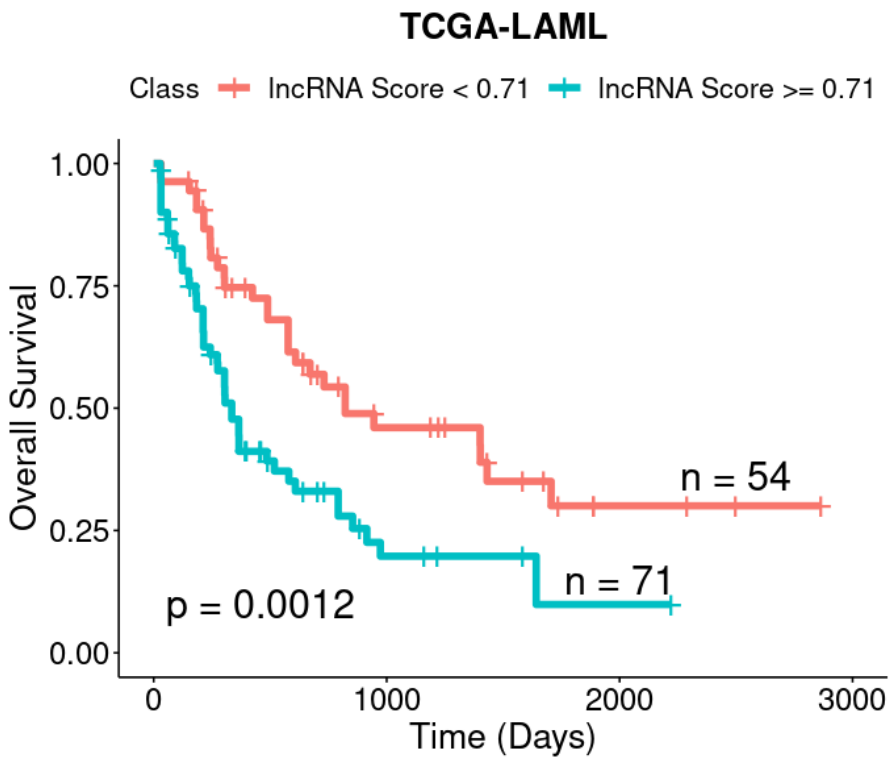
- Analysed the data in **identical manners** as the discovery cohort

- Observe whether the 10-lncRNA score retains prognostic significance on multi-variable analysis



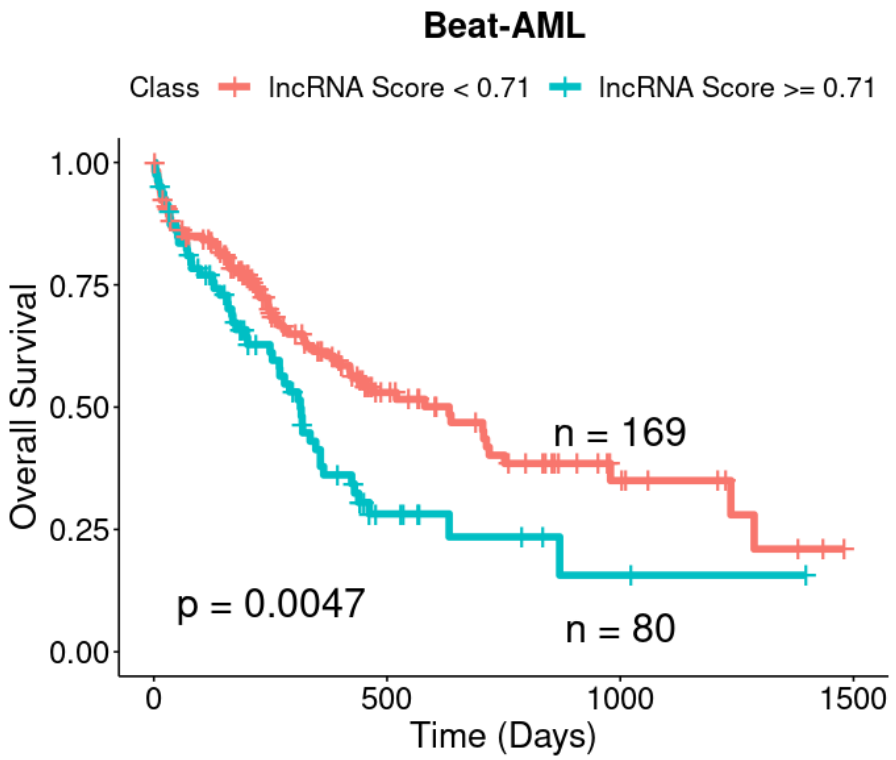
Validation of 10-lncRNA score

Multi-variable analysis in TCGA cohort



Validation of 10-lncRNA score

Multi-variable analysis in BeatAML cohort

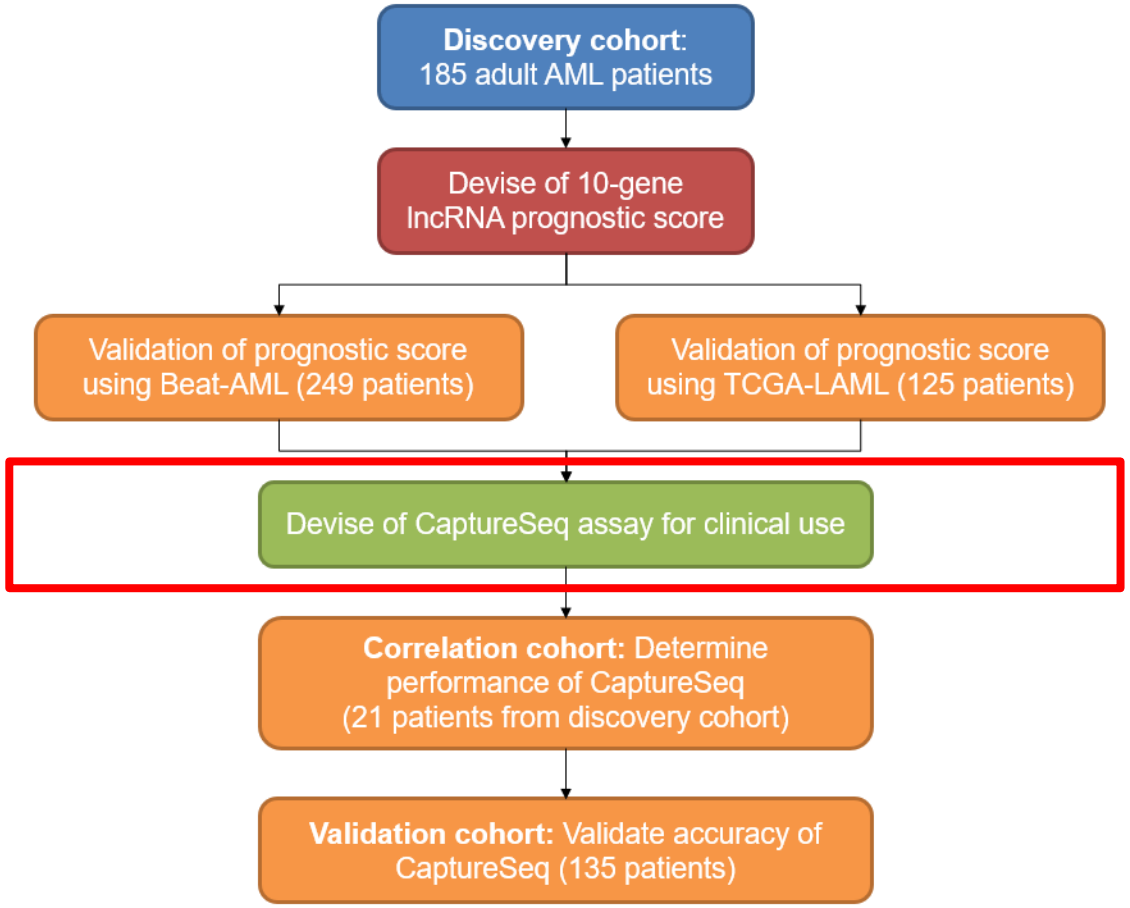


Clinical translation of research findings

Considerations for clinical applications

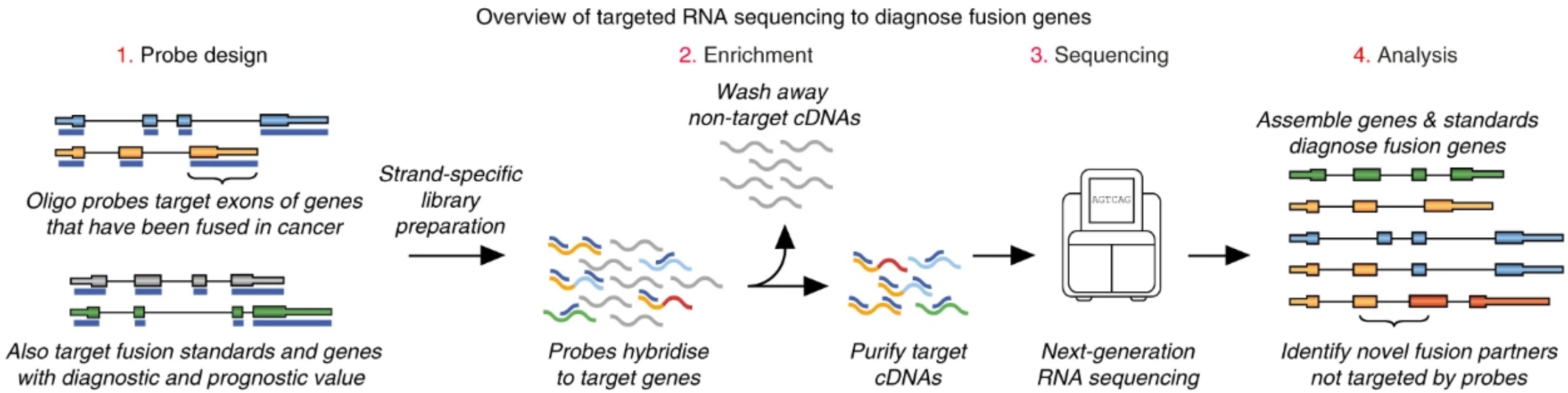
- Total transcriptome sequencing
  - Relatively expensive
  - Substantial portion of data may not have immediate clinical utility

- Enrichment of targeted genomic regions (CaptureSeq)
  - Cheaper
  - Targeted but more focused (more sequence reads to cover targeted regions)
  - Higher sensitivity



Clinical translation of research findings

Design of CaptureSeq panel for leukaemias

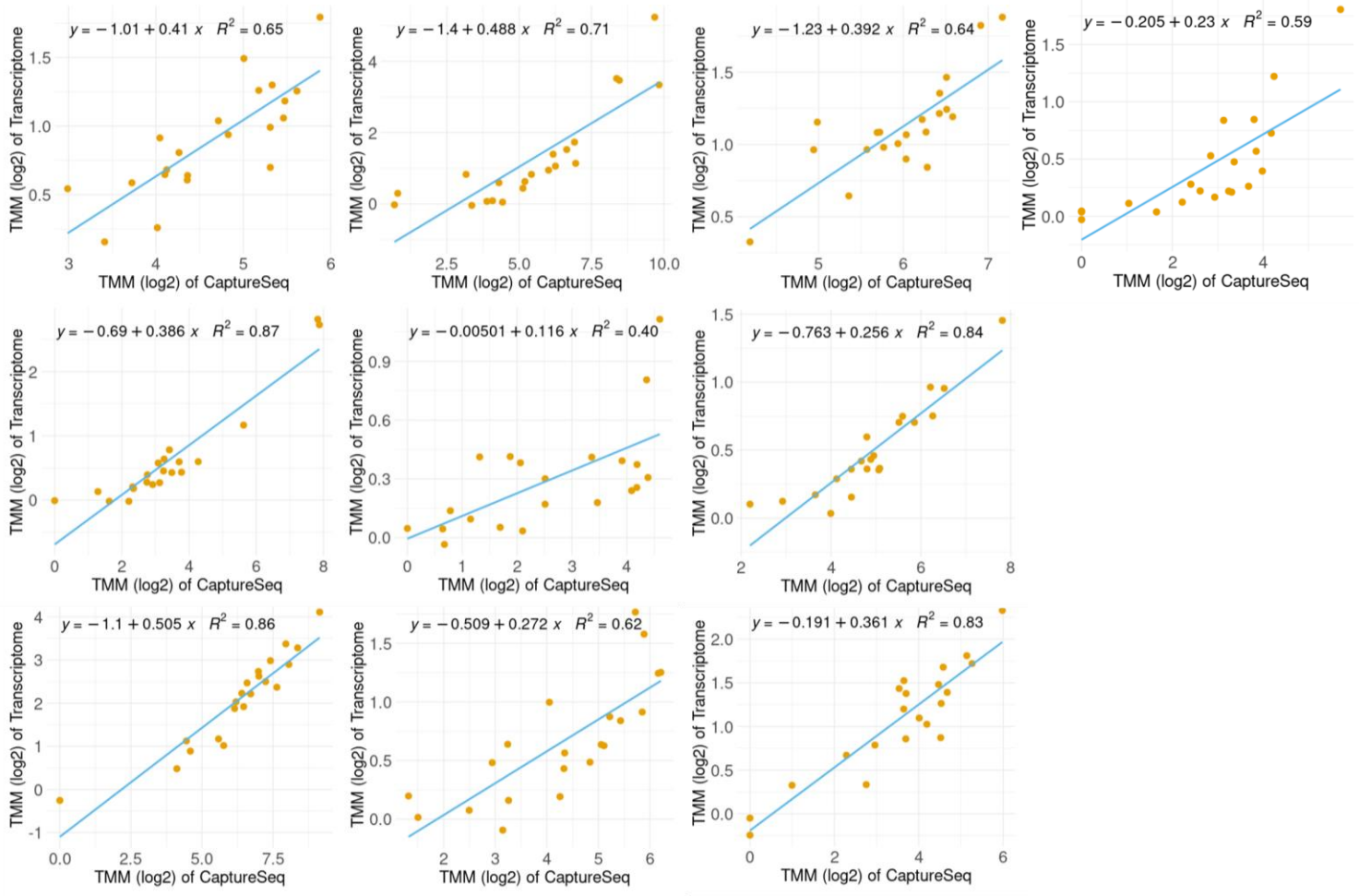


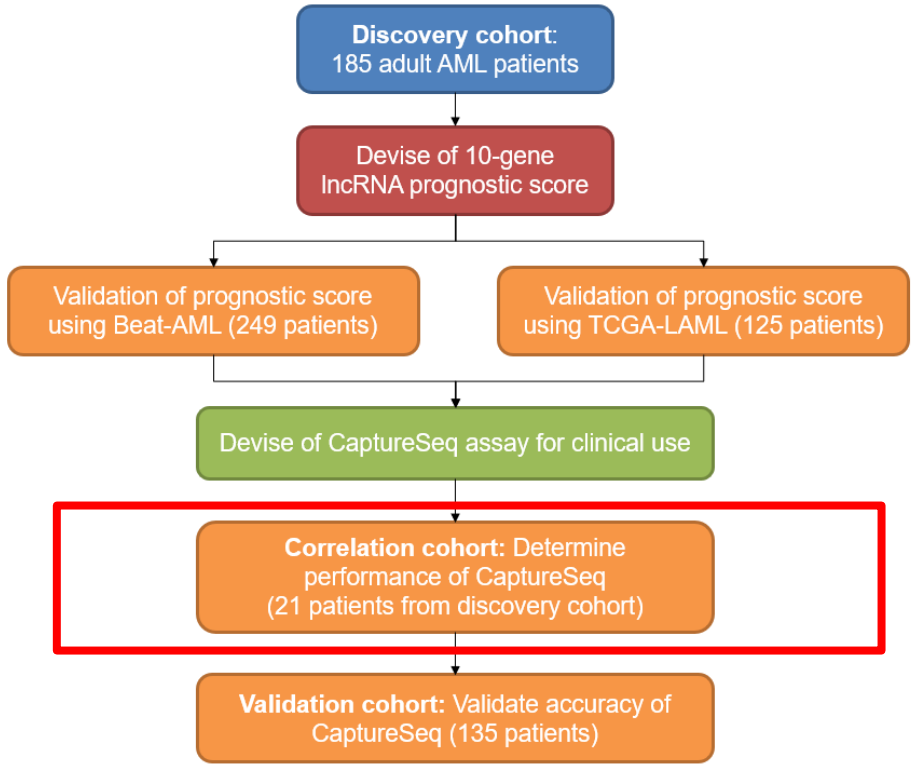
- Genes involved in **fusion** in leukaemias
- Genes for **expression profiling**: Coding genes and the 10 lncRNAs

Heyer EE et al. *Nat Commun* 2019;10:1388.

Performance of CaptureSeq panel

Correlation of 10 lncRNAs between transcriptome and CaptureSeq

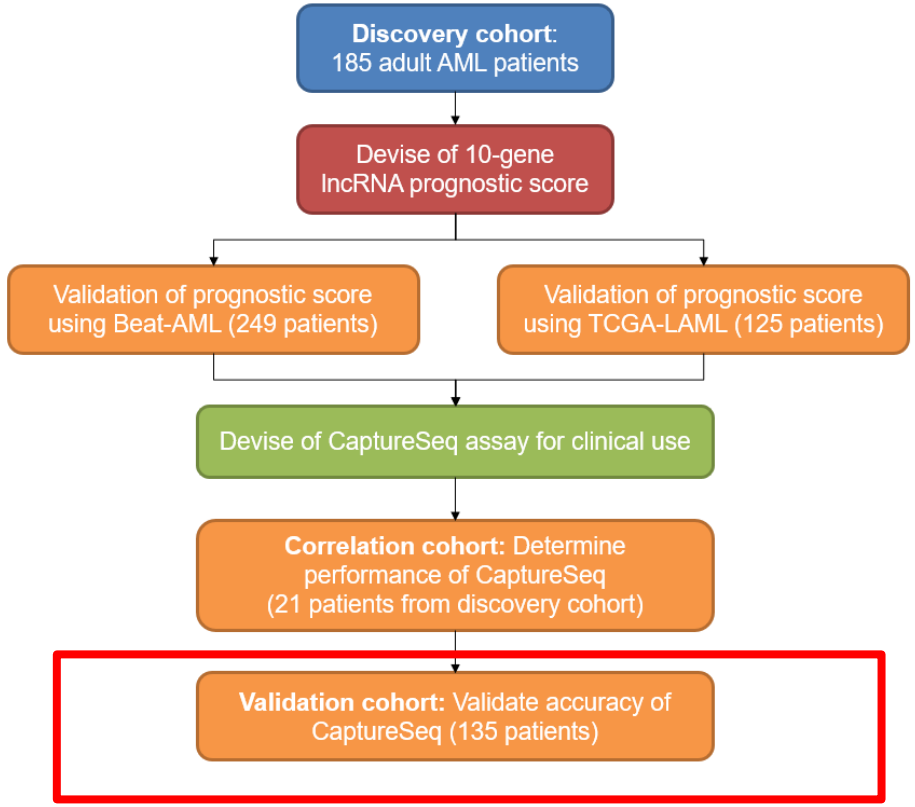




Validation of CaptureSeq panel

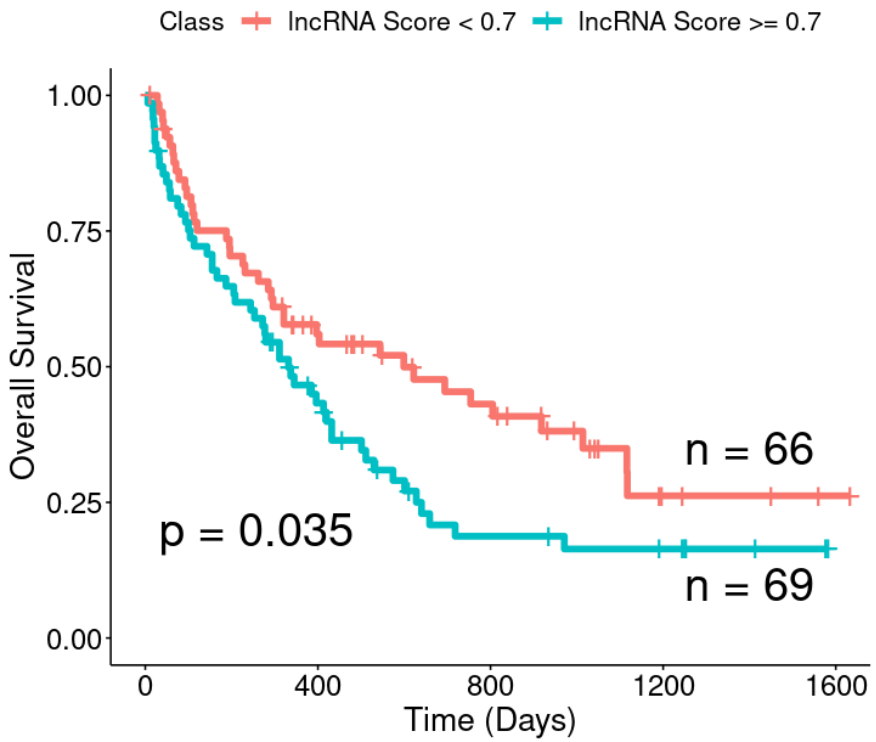
Independent prognostic evaluation in 135 patients

- Period: 2019-2022
- 135 consecutively recruited newly diagnosed AML patients
- Median follow-up: 335 days
- Comparison between validation cohort vs. discovery cohort
  - Older patients, e.g. 40% vs. 17% over age 70
  - More adverse risks patients, 50% vs. 38%



Validation of CaptureSeq panel

Independent prognostic evaluation in 135 patients

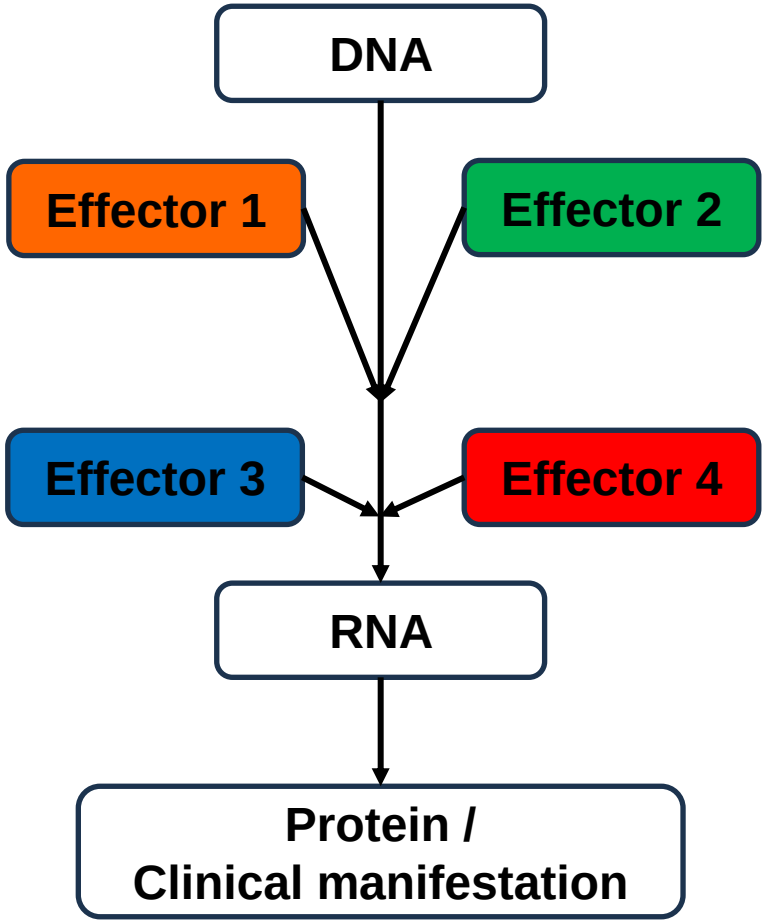


Project Summary

- Use of machine learning to identify a 10-lncRNA prognostic score
- The lncRNA prognostic score can reproducibly predict clinical outcomes of AML patients, with independent effects from the currently established prognostic parameters
- Rigorous validation of the lncRNA prognostic score using large public data sets
- CaptureSeq assay is devised for clinical translation and represents a viable option for refinement of established prognostic parameters in AML
- Substantiate the clinical utility of transcriptomics in informing the practice of precision medicine in leukaemia patients

Healthcare Implications of Project

- This project provides proof-of-concept that RNA sequencing contributes to unique information to inform clinical management
- **Genome (DNA) sequencing** provide 1st dimension of information
- **Transcriptome (RNA)** provides an indispensable 2nd dimension, i.e. functional genomics
- Cancer classification is incorporating transcriptomic information
- Prime time to strengthen our efforts in functional genomic investigations to harness **multi-omic information** to empower practice of precision medicine



Acknowledgement

- Cytogenetics and Genomics Laboratory, QMH
- Prof. Jason WH Wong, SBS, HKU
- Prof. SY Leung, SClinMed, HKU
- Prof. Anskar YH Leung, SClinMed, HKU
- Prof. H Sun, Chemical Pathology, CUHK

- Fundings
  - Health and Medical Research Fund, Health Bureau, HKSAR Government
  - Hong Kong Blood Cancer Foundation
  - Centre for Oncology and Immunology under the Health@InnoHK Initiative, funded by the ITC, HKSAR Government
  - Theme-based Research Scheme (T12-702/20-N), University Grants Committee, HK