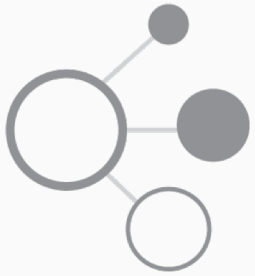


GRAPH+AI SUMMIT

organized by TigerGraph

Improving the Treatment of Acute Lymphoblastic Leukemia using Graph Analytics with AI and Machine Learning

Jesper Vang, PhD Student at Cancer Systems Biology group

DTU on Interregional Childhood Oncology Precision Medicine Exploration (iCOPE)

Jesper Vang, PhD Student at Cancer Systems Biology group

The rate of cancer in children rises in Europe by 6-10% per ten years, and mortality and treatment toxicity are still high.

To improve diagnostics, treatment, cure rates, and the overall life situation of children with cancer (at the hospital, in school, and at home).

How?

Build	Analyses	Ethical	Telepresence	e-health
Build a database and biobank for childhood cancer research based on ~2000 patients.	Carry out extensive germline and tumour DNA and transcriptomics (NGS) analyses to gain insights into cancer predisposition syndromes, somatic mutations, and tumor biology.	Address ethical issues arising from genome sequencing.	Develop telepresence robots to support children with staying connected to school during cancer treatment.	Investigate e-health as a tool to support healthcare at home.

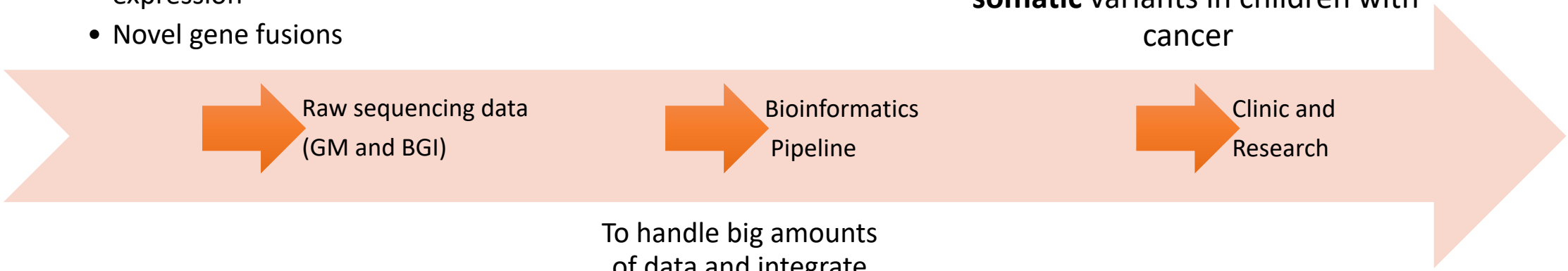
NGS Team at DTU:

Operational goal:

Types of data:

- **Germline DNA: Whole Genome Sequencing (WGS) data**
- **Tumour RNA: RNA-seq data**
  - Molecular subtypes from gene expression
  - Novel gene fusions

Research goals:
Germline-somatic continuum:
Find links between **germline** and **somatic** variants in children with cancer



To handle big amounts of data and integrate different data sources and extract biologically relevant signals through our pipelines

General workflow overview

Raw FASTQ (BGI and GM)

Sentieon® Pipeline :

Germline variation (WGS)

Tumour variation (RNAseq)

Somatic variation (WGS+RNAseq)

Tumour Gene expression

Gene Fusion

Structural Variant

SNV and short indel variant data

Data storage at Computerome:

Consistent data structure

Set of programmes to track and operate with data

Variant Database:

Variant annotation (Gene, transcript, freq. in the population GNOMAD)

Deleteriousness prediction (CADD, SIFT, PolyPhen, REVEL)

Allele Frequency in our cohort

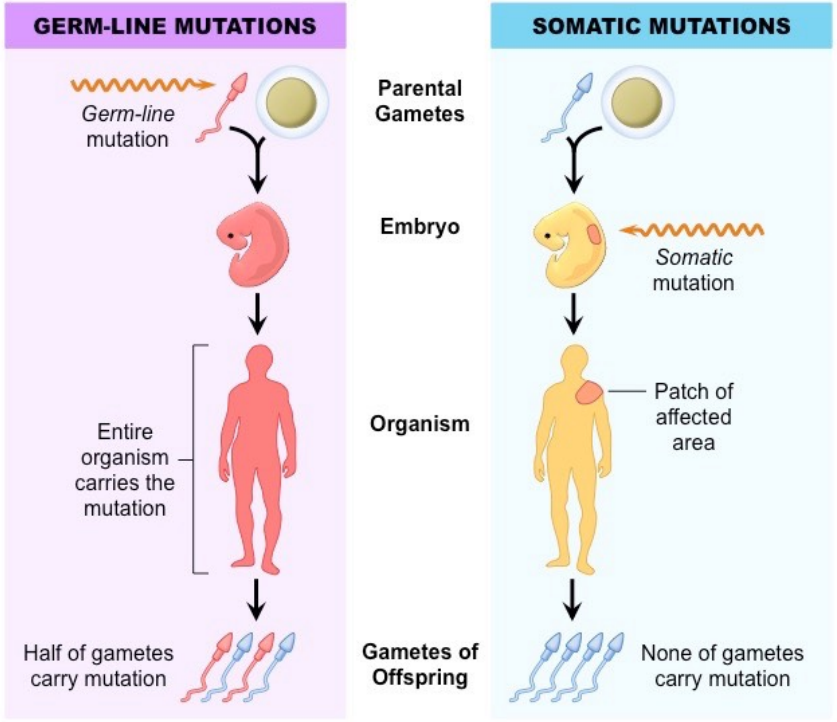
Germline Somatic 1:

Map and describe correlations of germline and somatic events in individual patients with gene expression and molecular signalling (Cancer Hallmarks, Cosmic Genes, DNA Repair genes), and their association to risk assessment and patient stratification

Study Cohort: 116 paediatric ALL patients

Data:

Whole Genome Sequencing for Germline DNA (N=116)
RNAseq for Tumour RNA (N=94)



Somatic Mutations
Present only in **some** cells
Not transmitted to offspring
Do **not fixate** in population

Germline Mutations
Present in **all** cells
Transmitted to offspring
Fixate in population (SNP)

Germline Somatic 2:

Normal Tissue (WGS)

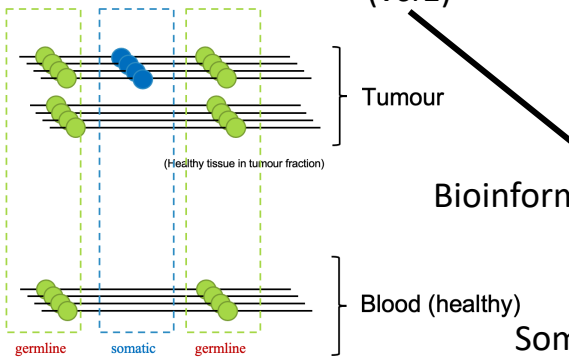
Tumour (RNAseq)

Bioinformatics pipeline

Bioinformatics pipeline

Germline Variation (vcf1)

Tumour Variation (vcf2)



Bioinformatics pipeline

Somatic Variation (vcf3)

DB

What are the consequences of the variation?

Evaluate Deleteriousness

CADD score
ENSEMBL annotation
Gnomad Pop. Freq
…

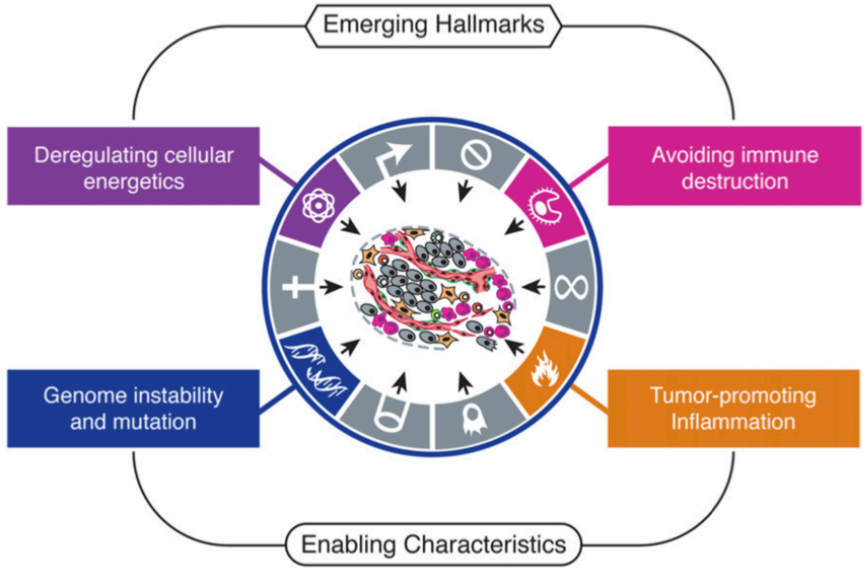
Germline predisposition genes
Tumour variation/genetics
Somatic variation

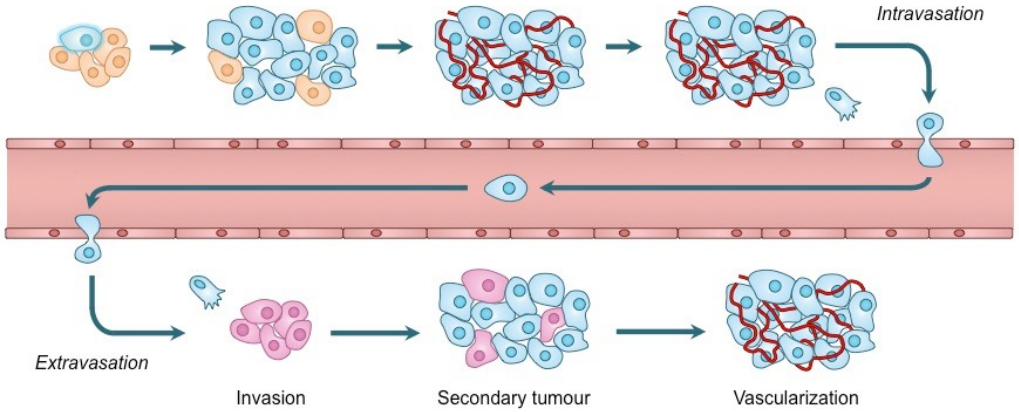
Is there any enrichment in the variation?

COSMIC (Catalog Of Somatic Mutations)
Cancer Hallmarks
DNA repair genes

Acquire functional capabilities

- Sustaining proliferative signalling
- Evading growth suppressors
- Resisting cell death
- Enabling replicative immortality
- Inducing angiogenesis
- Activating invasion and metastasis
- Emerging Hallmarks
- Enabling characteristics

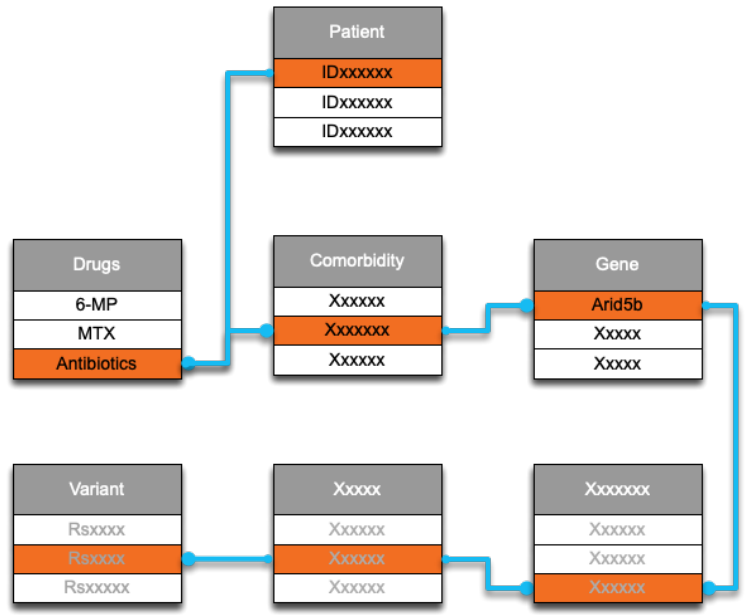


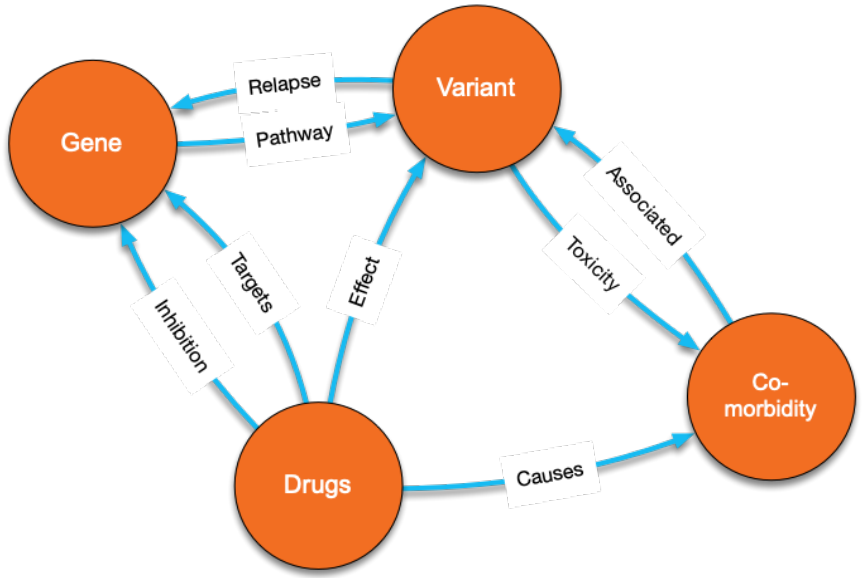


(ib.bioninja.com.au)

The hallmarks o cancer
Hanahan and Weinberg, Cell 2011

Why TigerGraph





Format	Nodes and edges	Tabels with rows and columns
Relationship	Considered data, represented by edges between nodes	Related across tables established using foreign keys between tables
Complex Queries	Run quickly and do not require joins	Require complex joins between tables

Thanks to the team:

Adrian Otamendi
PhD student
Done: June 2023

Olga Rigina
Data manager

Rikke L. Nielsen
Joint PostDoc
50% RH March 2023
50% DTU March 2021

Marianne Helenius
PhD student
Done: May 2023

Line Lund
Student assistant

Ramneek Gupta
10% Groupleader,
Associate Professor

Jonas V. Jensen
Student assistant

Frederik S. Gade
Student assistant

Jesper Vang
PhD student
Done: June 2022

Freja Hede
Student assistant

Elena Papaleo
Associate Professor

Carina T. Refsgaard
Master student

Bernadette Christiansen
Student assistant

Sara Garcia
PhD student
Done: Apr 2021

Anna S. Lassen
Student assistant

Kjeld Schmiegelow MD
Professor and
Senior Hospital Physician

Ida Hansen
Student assistant