9° WORKSHOP IN EMATOLOGIA TRASLAZIONALE

DELLA SOCIETÀ ITALIANA DI EMATOLOGIA SPERIMENTALE

Bologna, Aula "G. Prodi", 19-20 maggio 2025

Sequenziamento long-read: Principi e Applicazioni

Niccolò Bartalucci

University of Florence

Disclosures di Niccolò Bartalucci

Company name	Research support	Employee	Consultant	Stockholder	Speakers bureau	Advisory board	Other
none	none	none	none	none	none	none	none

The Next Generation Sequencing

1888
Discovery of Chromosomes

1953
Discovery of DNA's double helix

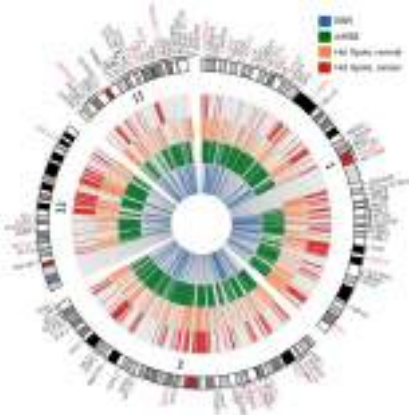
1977
Publication of Sanger method

1983
Developing of PCR

1990
Launch of Human Genome Project

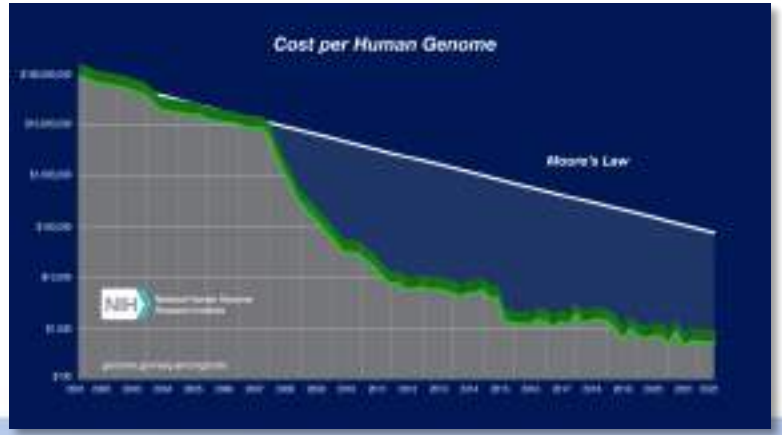
2003
Completion of Human Genome Project

2005
First NGS sequencer



Cisneros L et al. PlosOne, 2017

Trudsø LC et al. PlosOne, 2020



1888
Discovery of Chromosomes

1953
Discovery of DNA's double helix

1977
Publication of Sanger method

1983
Developing of PCR

1990
Launch of Human Genome Project

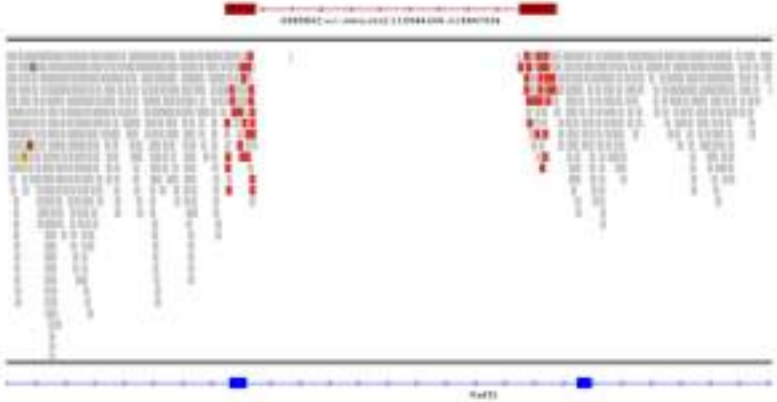
2003
Completion of Human Genome Project

2005
First NGS sequencer

2011
First Long-Read Sequencer

The Third Generation Sequencing

Humans differ for >3000 deletions sizing >500bp



~5% of human genome is duplicated

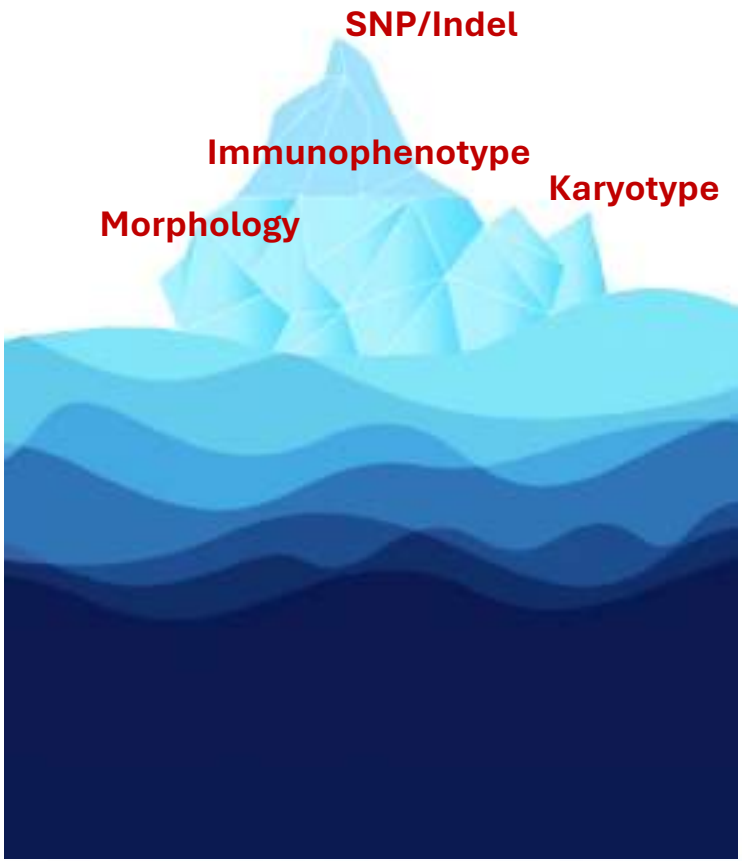


Table 2. Gene mutations in myeloid neoplasms and leukemia indicated for clinical testing

Indication	Single gene mutations	Structural variants
MDS, MDS/MPN, cytopenia	ASXL1, CALR, CBL, CSF3R, DNMT3A, EZH2, IDH1, IDH2, JAK2, KIT, KRAS, MPL, NRAS, PTPN11, RUNX1, SF3B1, SRSF2, STAG2, TET2, TP53, U2AF1, WT1, ZRSR2	
MPN and myelodysplasia	ASXL1, CALR, CBL, CSF3R, DNMT3A, EZH2, IDH1, IDH2, JAK2, KIT, KRAS, MPL, NRAS, PTPN11, RUNX1, SF3B1, SRSF2, STAG2, TET2, U2AF1, ZRSR2	BCR::ABL1
Eosinophilia	ASXL1, CBL, DNMT3A, EZH2, KRAS, NRAS, RUNX1, SF3B1, SRSF2, STAG2, TET2, U2AF1	BCR::ABL1, FGFR1, IG::FGFR1, JAK2::PCM1, PDGFRA, PDGFRB
AML	**Genes required for diagnosis and risk stratification:** ASXL1, BCOR, CEBPA, DDX41, EZH2, FLT3-ITD, FLT3-TKD, IDH1, IDH2, KMT2A-PTD, NPM1, RUNX1, SF3B1, SRSF2, STAG2, TP53, U2AF1, ZRSR2 **Additional genes recommended to test for at diagnosis and for use in disease monitoring:** GATA2, MECOM(EVI1), SH2B3, SETBP1, SF3B1, PHF6, PPM1D, PTPN11, RAD21, SETBP1, TERT, WT1	BCR::ABL1, CBFB::MYH11, DEK::NUP214, KMT2A::rearranged, MECOM::rearranged, NUP98::rearranged, RBM15::MRTFA, RUNX1::RUNX1T1
B-ALL	CRLF2, EBF1, ETV6, IDH1, IDH2, IKZF1, PAX5, PTPN11, SETD2, SH2B3, TP53	ABL1::rearranged, BCR::ABL1, CDX2::rearranged, CEBPA::rearranged, DUX4::rearranged, ETV6::ABL1, ETV6::RUNX1, HLF::rearranged, KMT2A::rearranged, MEF2D::rearranged, NUTM1::rearranged, PAX5::rearranged, PAX5::ETV6, PAX5::JAK2, PAX5::P80R, TCF3::HLF, ZNF384::rearranged
T-ALL	DNMT3A, ETV6, EZH2, FBXW7, FLT3, IDH1, IDH2, IL7R, JAK1, JAK3, KRAS, NRAS, NOTCH1, NRAS, PHF6, PTEN, U2AF1, WT1	BCL11B::rearranged, KMT2A::rearranged, MYB::NFIB, NUP214::ABL1, NUP214::DEK, TAL1::rearranged, TLX1::rearranged, TLX3::rearranged

Duncavage EJ et al. Blood, 2022

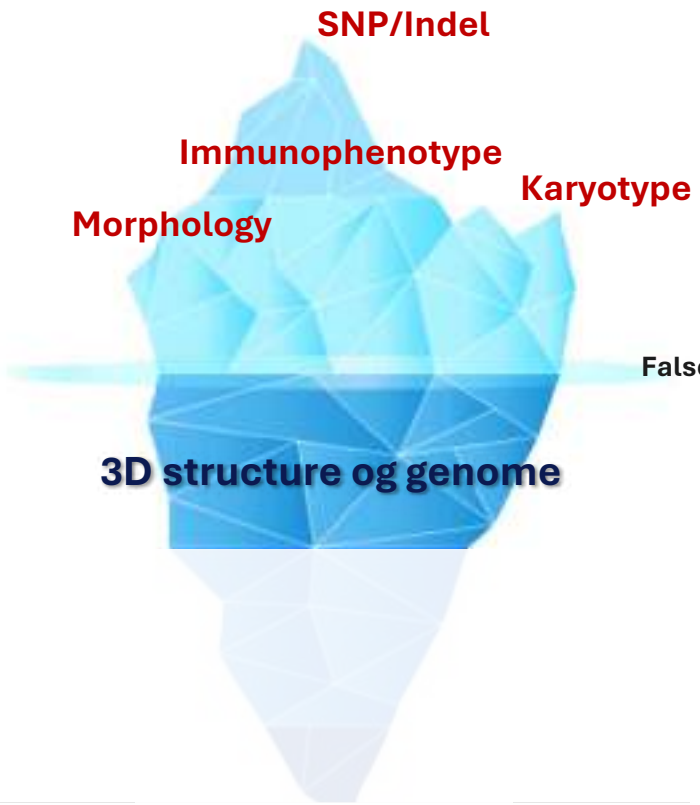
Table 6. 2022 ELN risk classification by genetics at initial diagnosis*

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§§

Dohner H et al. Blood, 2022

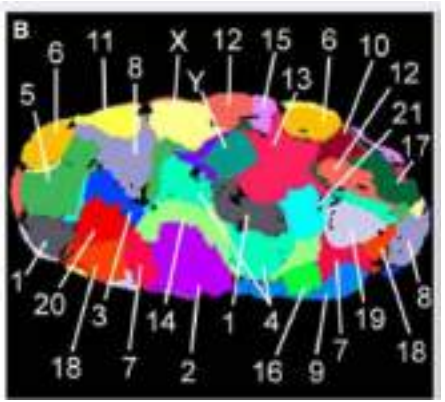
Greenfield G et al. J Hematol Oncol, 2021

Three-Dimensional Genome Architecture

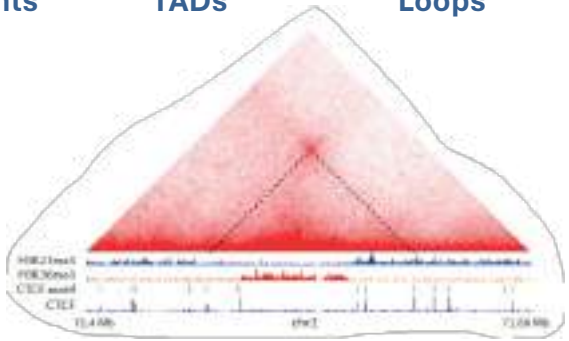


Chromosome Territories | Compartments A/B | TADs | Loops

False color representation of CTs in mid-section classified by goldFISH



Bolzer A et al. Plos Biol, 2005

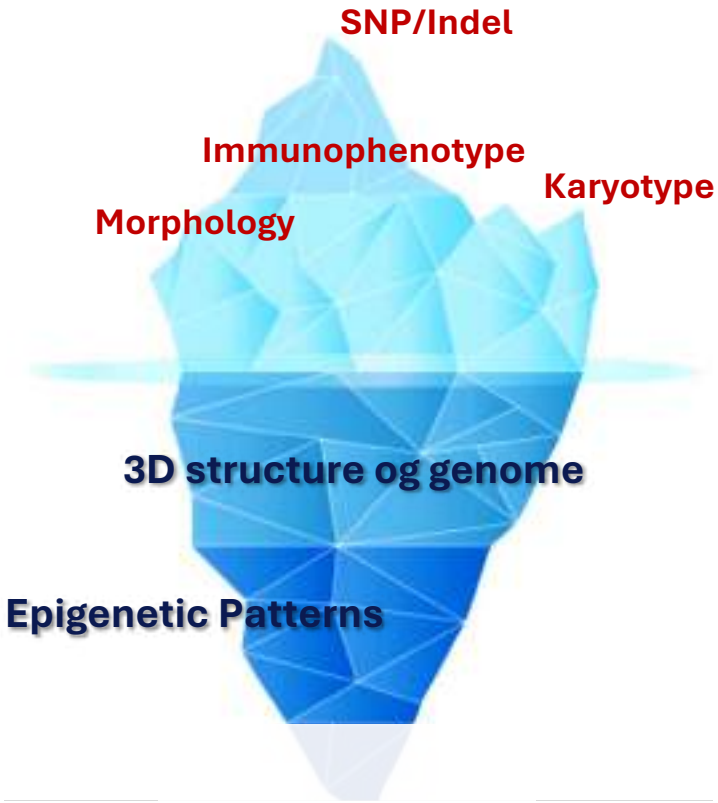


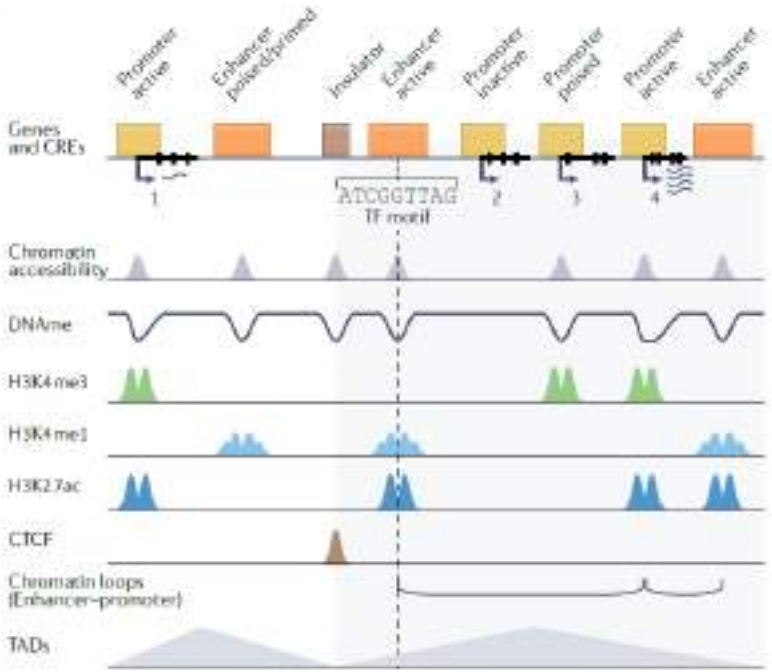
REGULAR ARTICLE

blood advances

AML with complex karyotype: extreme genomic complexity revealed by combined long-read sequencing and Hi-C technology

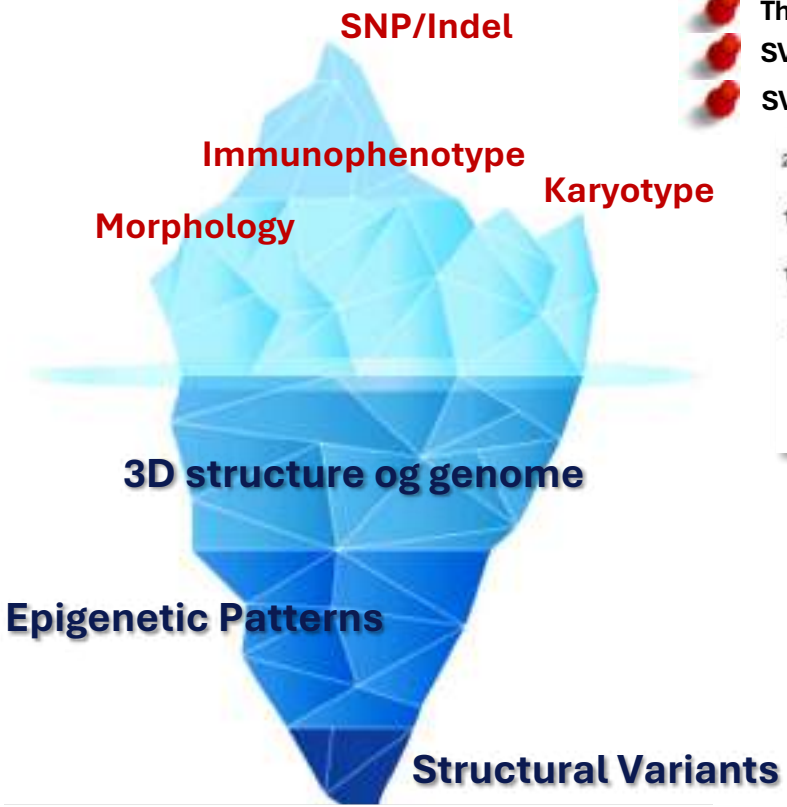
Epigenetic Regulation



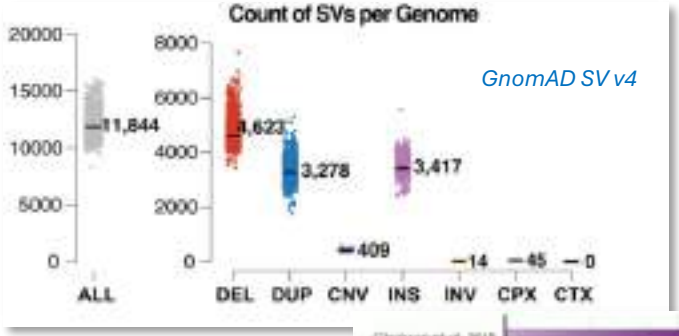


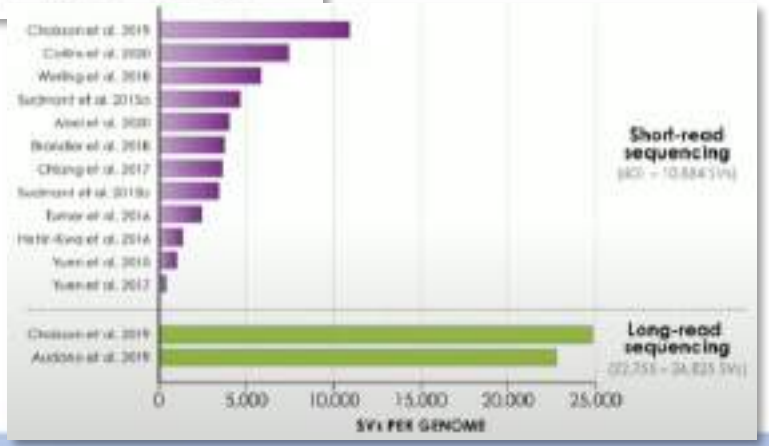
Preissl S et al. Nature Rev Genet, 2023

Structural Variants (SV)



- **The operational range of SVs includes events > 50bp.** *Alkan C, Nat Rev Genet, 2011*
- **SVs contribute to germline and somatic diseases.** *Talkowski M, Cell, 2012*
- **SVs are complex and important driver of human evolution.** *Soto DC, Am J Biol Anthropol, 2023*





Long-Read Sequencing (LRS)

- Long-range Sequencing
- Real-Time Analysis
- Direct Sequencing

DIAGNOSTICS

Acute Leukemia Classification Using Transcriptional Profiles From Low-Cost Nanopore mRNA Sequencing

PLOS GLOBAL PUBLIC HEALTH

Rapid detection of myeloid neoplasm fusions using single-molecule long-read sequencing

Review Article | Published: 28 March 2025

Computational analysis of DNA methylation from long-read sequencing

Yilei Fu, Winston Timp & Fritz J. Sedlazeck

Nature Reviews Genetics (2025) | Cite this article

Cell Press — Trends in Genetics

Review

A blood drop through the pore: nanopore sequencing in hematology

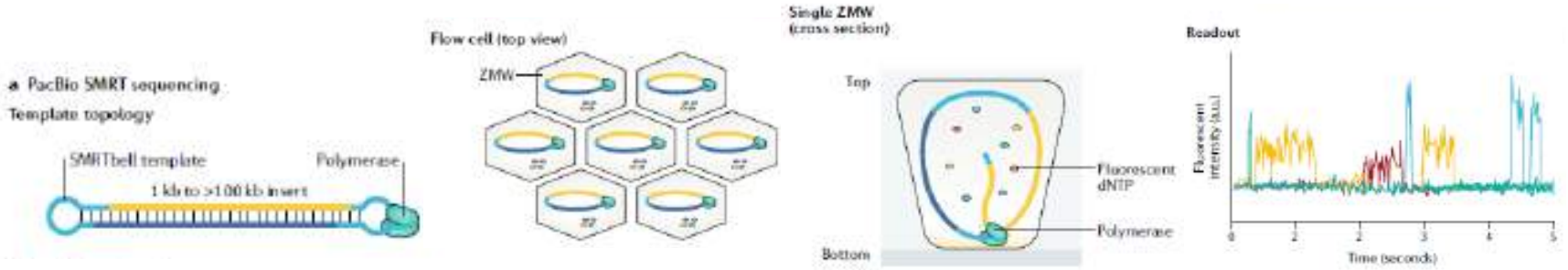
Migrating to Long-Read Sequencing for Clinical Routine *BCR-ABL1* TKI Resistance Mutation Screening

Wesley Schaal, Adam Arneur, Ulla Olsson-Strömberg, Monica Hermanson, Lucia Cavelier and Ola Spjuth

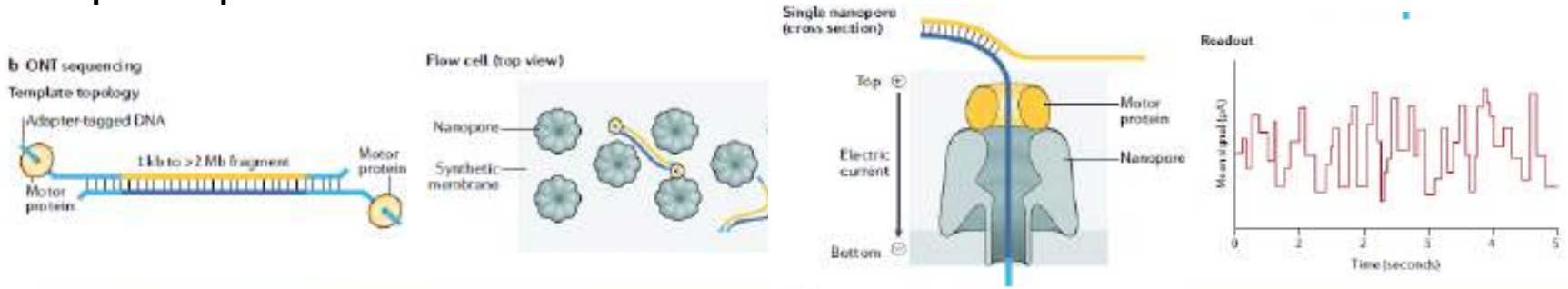
SAGE

LRS technologies

Single-Molecule Real-Time Seq



Nanopore Seq

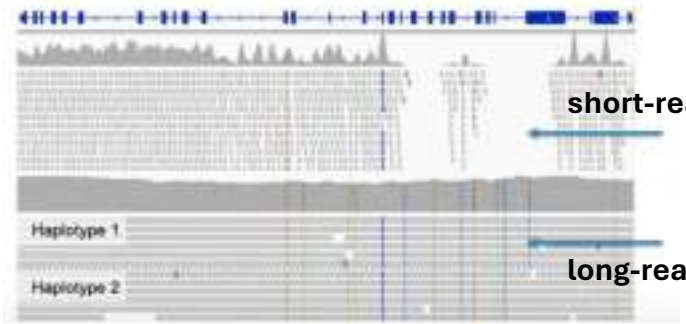


Whole Genome Sequencing (WGS)

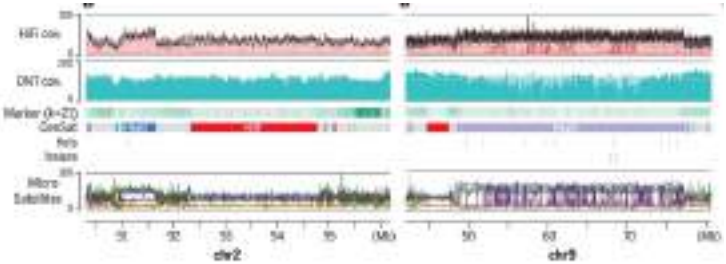
E. Coli genome: 4.6x10^6 bp

– 50 base read
– 92,000 "pieces"

– 500,000 base read
– 9 "pieces"



Telomere-to-Telomere (T2T) CHM13 release:
3.055 billion bp with gapless assemblies for all chr (-Y)

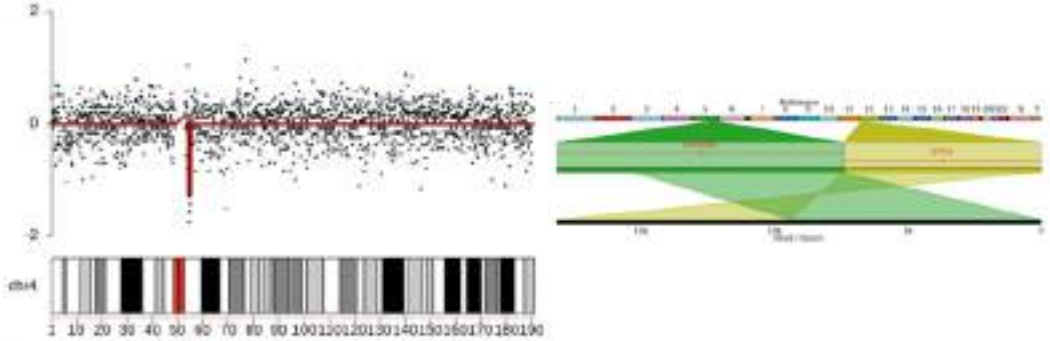


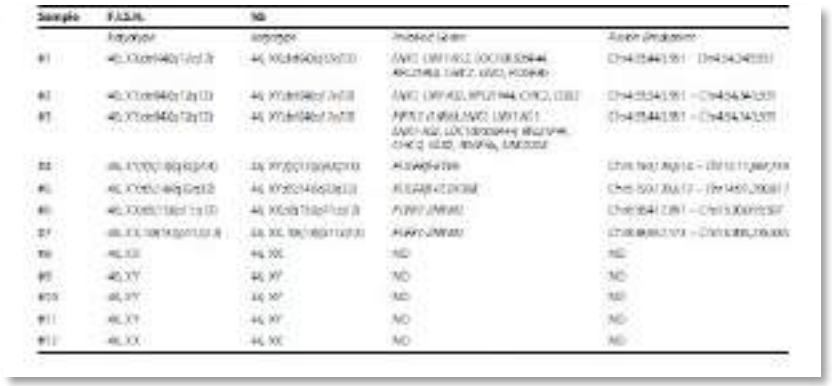
Nurk S et al. Science, 2022

STATISTICS	GRCH38	T2T-CHM13	DIFFERENCE (+%)
Segmental duplications			
Percentage of segmental duplications (%)	5.00	6.61	
Segmental duplication bases (Mbp)	150.71	201.93	+33.1
Number of segmental duplications	24097	40529	+72.3
RepeatMasker			
Percentage of repeats (%)	51.89	53.94	
Repeat bases (Mbp)	1,516.57	1,647.82	+8.7
Long interspersed nuclear elements	626.33	631.64	+0.8
Short interspersed nuclear elements	386.48	390.27	+1.0
Long terminal repeats	267.52	269.91	+0.9
Satellite	76.51	150.42	+96.6
DNA	108.53	109.35	+0.8
Simple repeat	36.5	77.69	+112.9
Low complexity	6.36	6.44	+4.6
Retroposon	4.51	4.65	+3.3
rRNA	0.21	1.71	+730.4

Detection of Gene Fusions

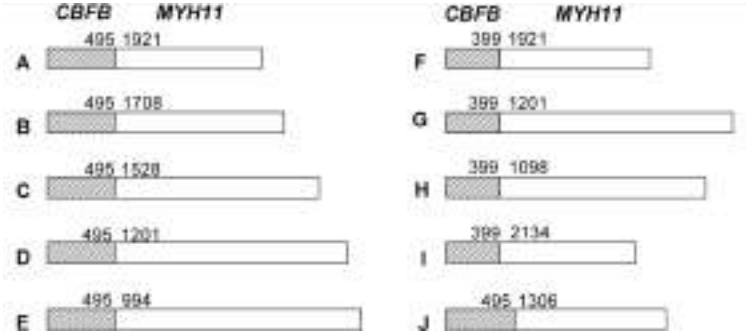
PDGFRα, *PDGFRβ*, *FGFR1*, *PCM1*::*JAK2* characterization



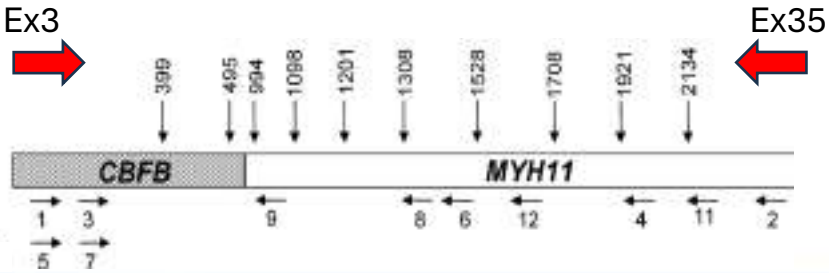


Romagnoli S et al. Biomark Res, 2021

CBFB::*MYH11* calling

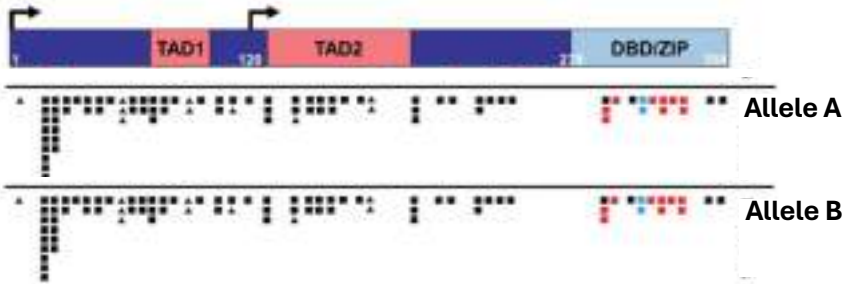


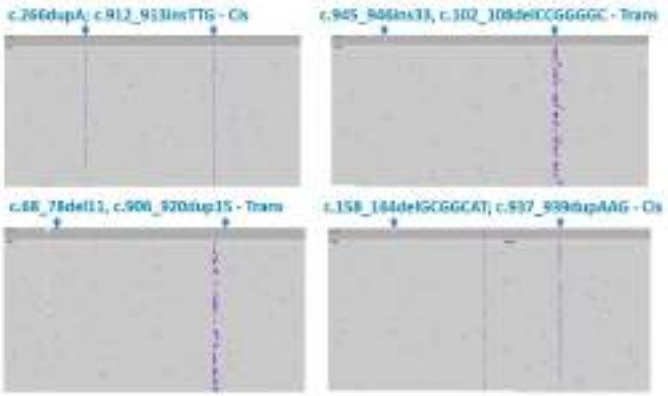
Monmia S et al. Leukemia Res, 2007



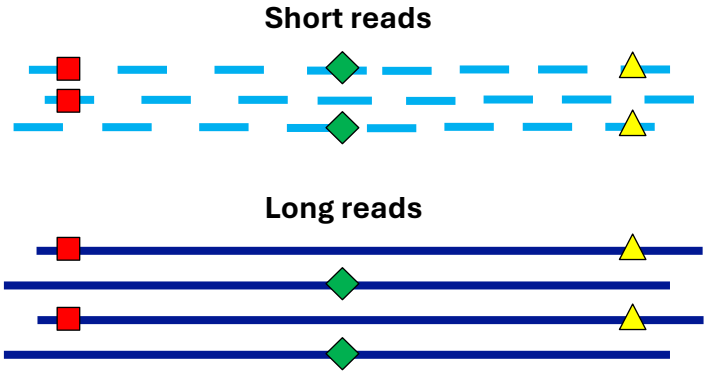
Variant Phasing

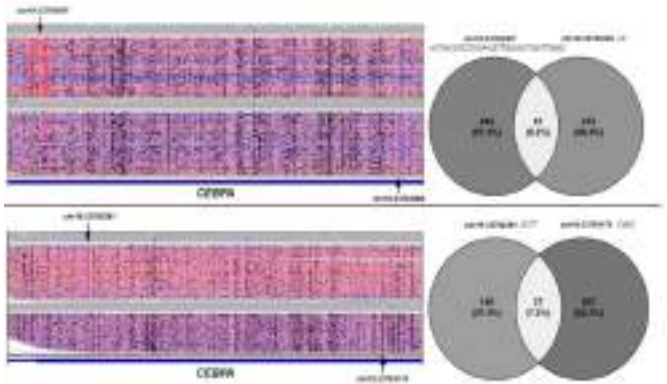
CEBPA mut phasing





Cai L C et al. poster Ass Mol Path, 2020





Cumbo C et al. Genes, 2019

Variant Phasing

nature communications

Article

Germline-somatic *JAK2* interactions are associated with clonal expansion in myelofibrosis

Received: 7 January 2022

Accepted: 26 August 2022

Published online: 08 September 2022

Check for updates

Derek W. Bowen[1,2], Weiyin Zhou[1,2], Youjin Wang[1], Kristine Jones[1,2], Wen Luo[1,2], Casey Dagnall[1,2], Kodect Teshome[1,2], Alyssa Klein[1], Tongwu Zhang[1], Shu-Hong Lin[3], Olivia W. Lee[3], Sainab Khan[3], Jacqueline S. Vo[1,2], Amy Hutchinson[1,2], Jia Liu[1,2], Jiahui Wang[1,2], Bin Zhu[1,2], Belinda Hicks[1,2], Andrew St. Martin[4], Stephen R. Spellman[4], Tao Wang[1,2], H. Joachim Deeg[3], Vikas Gupta[5], Stephanie J. Lee[1,2], Neal D. Freedman[1], Meredith Yeager[1,2], Stephen J. Chanock[1], Sharon A. Savage[1], Wael Saber[6], Shahrukh H. Gadiwala[1,2] & Mitchell J. Machiela[1,2]

Supplemental Table 5. Germline haplotypes by *JAK2*^{V617F} mutation status

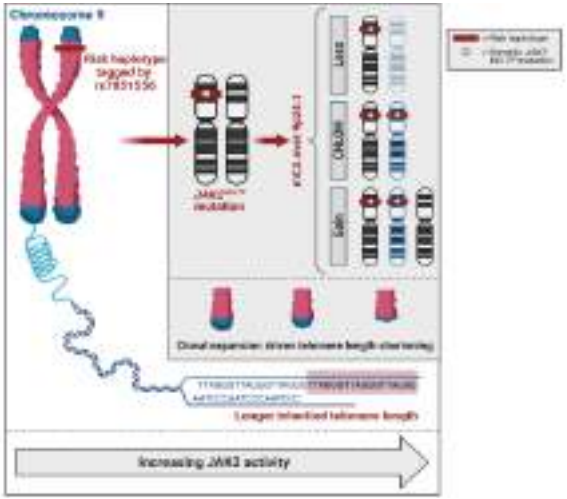
	JAK2^{V617F} mutation status			
	Mutation	No Mutation	Total	p-value[a]
N	562	1,286	1,848	
Germline Haplotype				
GGC	370 (65.84)	384 (28.86)	754	1.23×10^{-26}
TCT	187 (33.27)	887 (68.97)	1074	2.25×10^{-22}
GCC	4 (0.71)	6 (0.47)	10	0.5044
TGC	1 (0.18)	6 (0.47)	7	0.6831
CCT	0 (0.00)	1 (0.08)	1	1
GCT	0 (0.00)	1 (0.08)	1	1
TCC	0 (0.00)	1 (0.08)	1	1

[a]Two-sided binomial test

GGC= Germline risk haplotype

N= Number of germline haplotypes

- MF with high-frequency JAK2 mCAs have marked reductions in telomere length suggesting a relationship between telomere biology and clonal expansion.
- Germline variation at the 9p24.1 risk haplotype confers elevated risk of acquiring *JAK2* ^{V617F} mutations
- 370/562 (65.8%) patients had *JAK2* ^{V617F} mutation in cis with the risk haplotype



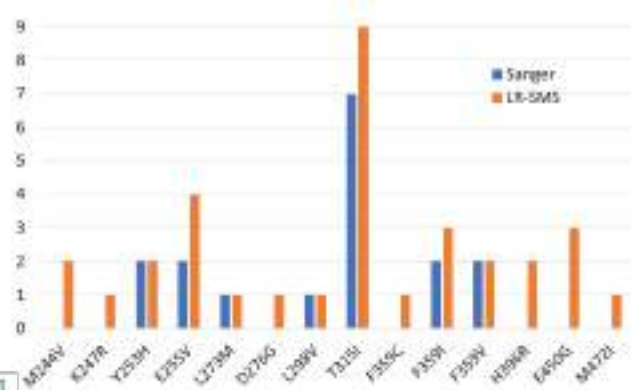
Rapid Time-to-Results

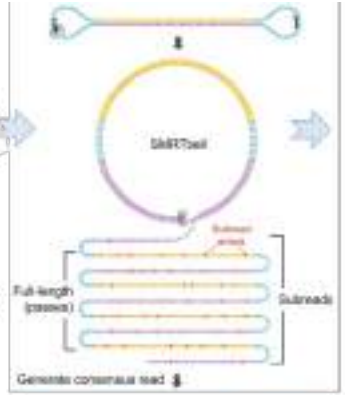
Migrating to Long-Read Sequencing for Clinical Routine *BCR-ABL1* TKI Resistance Mutation Screening

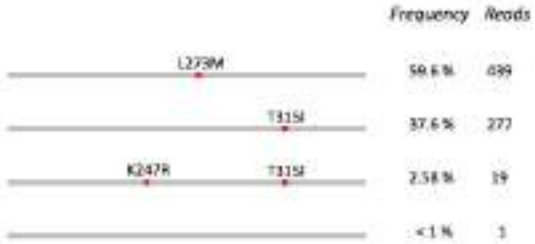
Wesley Schaal[1,2], Adam Ameur[2,3], Ulla Olsson-Strömberg[4], Monica Hermanson[3], Lucia Cavelier[3] and Ola Spjuth[1,2]

[1]Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden. [2]Pircor Bio AB, Uppsala, Sweden. [3]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [4]Department of Medical Sciences, Uppsala University Hospital, Uppsala, Sweden.

Cancer Informatics
Volume 21: 1–8
© The Author(s) 2022
DOI: 10.1177/11769351221110872
SAGE







2h

Rapid Time-to-Results

Sample → RNA → *FLT3* exon 13-16 cDNA → PCR and Adapter ligation → Load & Run → Live Alignment → Live ITD-calling

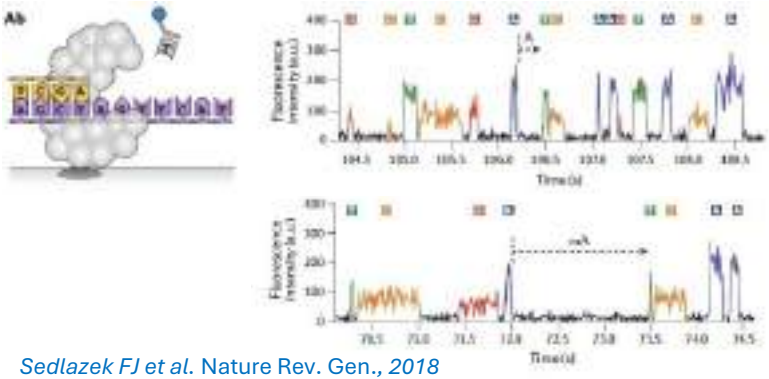
2h | 2h | 1'-3h

live analysis

ID	CHROM	POS	ALT SEQUENCE	SV TYPE	SV LENGTH	DR	DV	TIME
12	chr13	28034134	GAGATCATATTCATA	INS	18	13537	2971	7h
13	chr13	28034132	ATTGTAAGGCTACCAAACTCTAAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATAT	INS	63	26122	25	8h
18	chr13	28034132	AAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATAT	INS	42	129125	1252	4h45'
25	chr13	28034130	CCCATTTGAGATCATATTCAT	INS	21	22854	2397	6h45'
8	chr13	28034176	AGGTGGTAAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATATTCTCTGAAATCAACGTAGAAG TACTCATTATCTGAGGAGCCGGT	INS	93	117411	338	7h30'
16	chr13	28034154	GATCATATTCATATTCTCTGAAATCAACGTAGAAGT	INS	36	5772	549	7h
52	chr13	28034158	TTCATATTCTCTGAAATCAACGTAGAAGTACTC	INS	33	4990	562	6h15'
54	chr13	28034133	CTCTTGAAATCAGGGATTCATATT	INS	24	30342	3930	4h30'
59	chr13	28034181	GGGGGCCTAAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATATTTCTTGAAATCAACGTAGAA GTACTCATTCTGTAAAAGTATCAGTCACCT	INS	99	139003	167	8h
61	chr13	28034138	AAAAAACTCCATTTGAGATCATATTCATATTCTCTG	INS	36	36898	254	5h
39	chr13	28034137	AATGAGATCATATTCATATTCTCT	INS	24	3189	966	6h

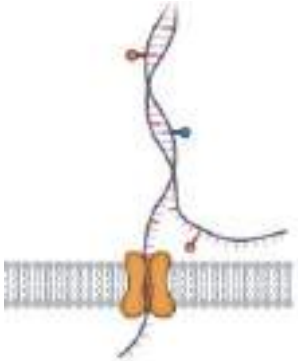
LRS Direct Methylation Calling

SMRT sequencing



Sedlazek FJ et al. Nature Rev. Gen., *2018*

Nanopore Sequencing



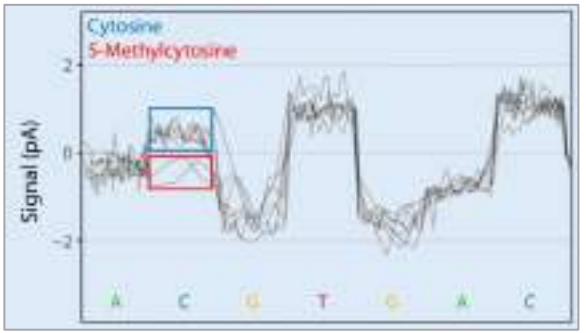


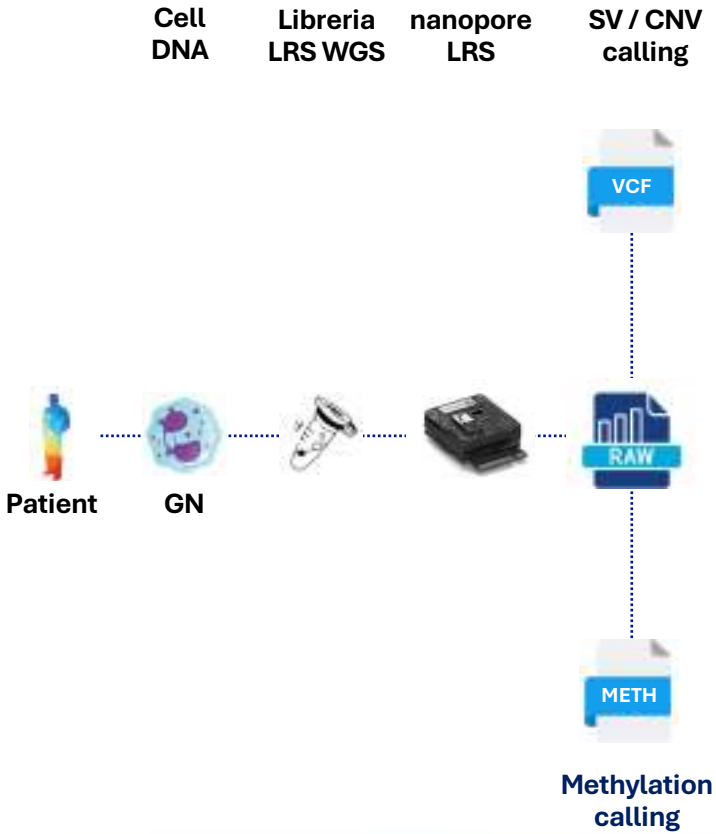
Table 1 | Comparison of methods for profiling methylation

Platform	Cost per sample (USD)*	Scale	Library preparation	DNA input	Read accuracy (Q20+) (%)	SmC calling accuracy (%)
Microarray	~400	935,000 CpGs	Chemical treatment needed	Low with multiple samples per array	99.75	95–99
Short-read, treatment	~850	28 million CpGs (whole genome)		Low		
ONT long read	~2,000		Direct readout	Medium	97–99	99
PacBio long read				High, high molecular weight preparation	>99.5	90

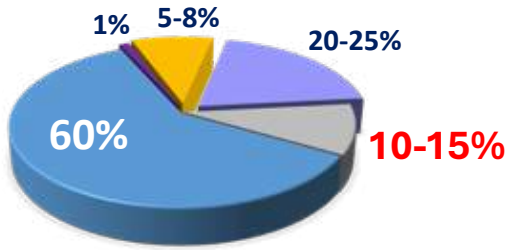
Fu Y et al. Nature Rev. Gen., 2025

Umair M et al. Nature Comm., 2023

Simultaneous SV and Methylation profiling



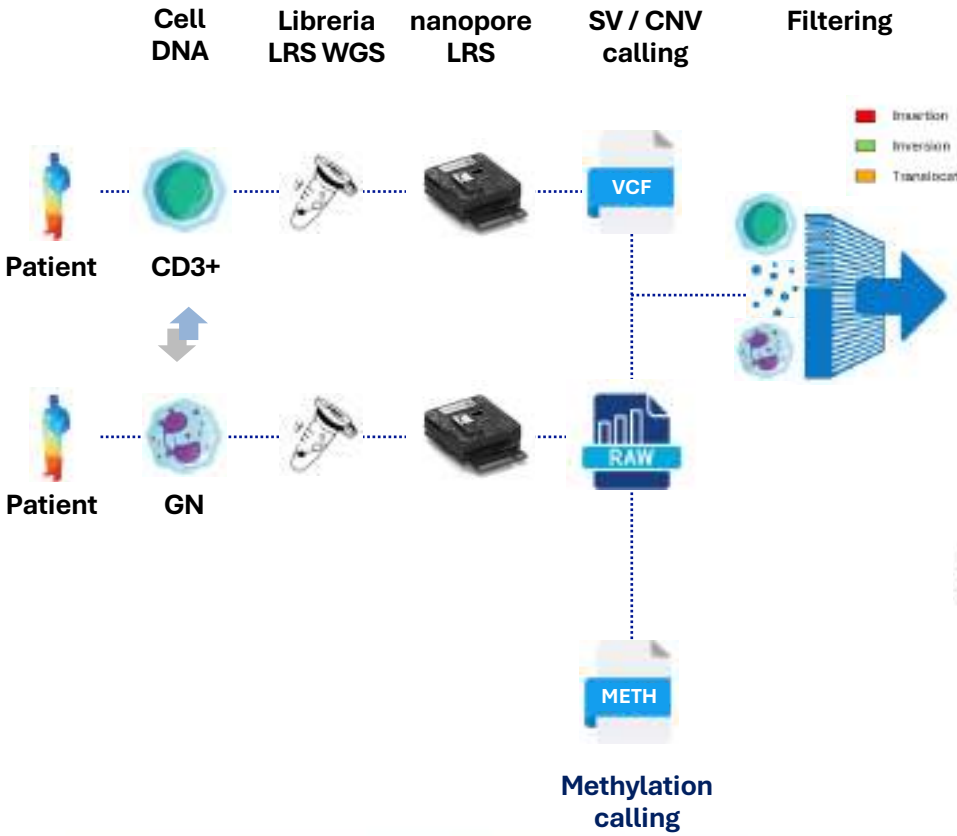
Primary Myelofibrosis



- *JAK2* V617F
- *CALR* T1/T2-like
- *MPL* W515*
- Non-canonical *MPL* or *JAK2*
- **Unknown «Triple Negative»**

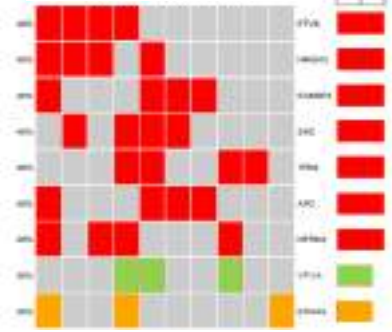
Milosevic JD et al. Blood, *2016*

Simultaneous SV and Methylation profiling

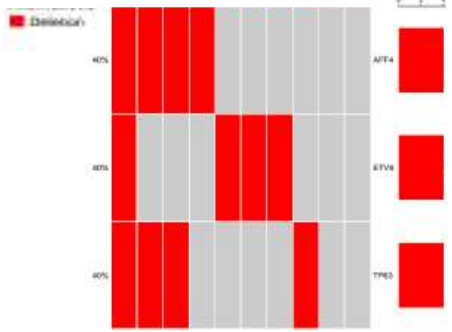


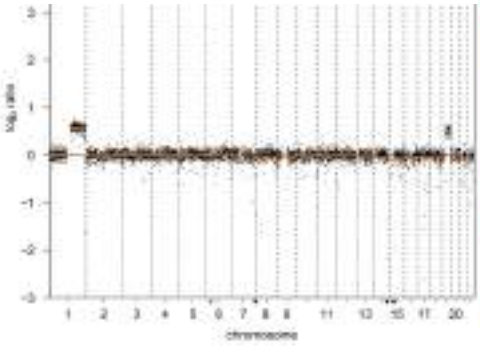
Somatic SV / CNV

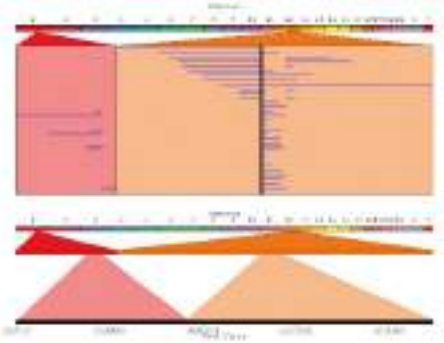
Cancer Gene Census



Network of cancer genes

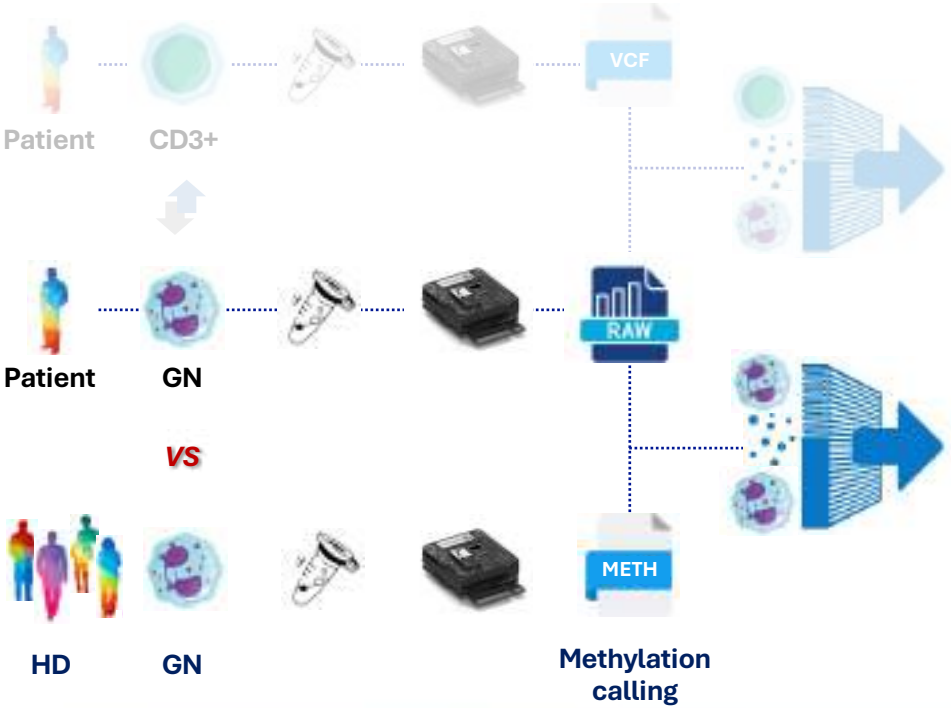


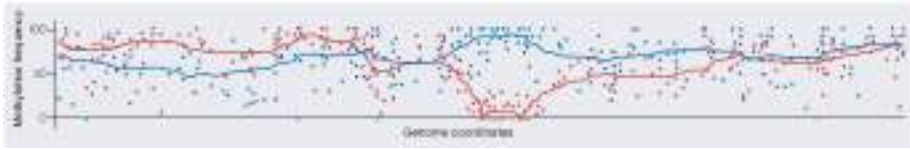


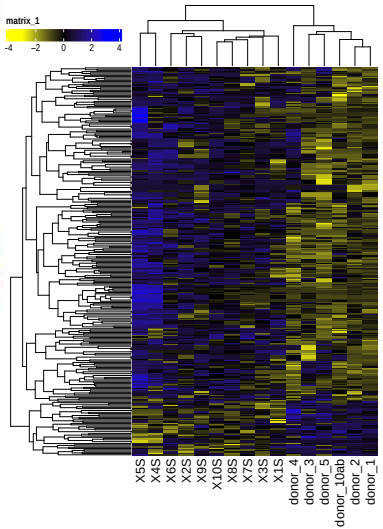


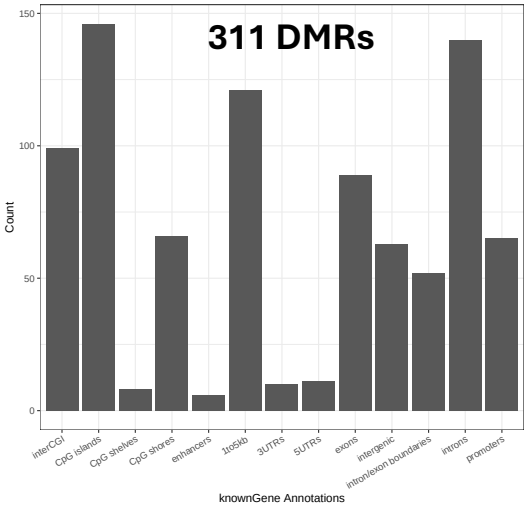
Simultaneous SV and Methylation profiling

Cell DNA | Libreria LRS WGS | nanopore LRS | SV / CNV calling | Filtering | Differentially Methylated Regions (DMRs)



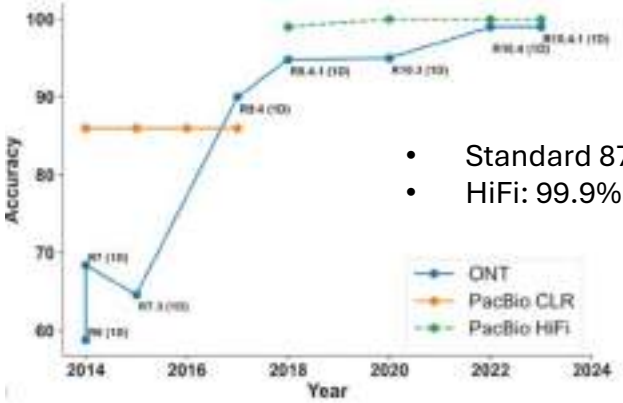






Major Challenges of LRS

Error rate



- Standard 87-98% reads Q20
- HiFi: 99.9% reads Q20

Annotation

- Few public dataset
- Pop SV repostory (GnomAD SV v4)
- Mapping

- Read N50 of 54 kbp
- Mean coverage 37x
- ~24,500 SVs per genome

Gustafson JA et al. Genome Res., 2024

Sample requirements

- Amount of sample (1ug)
- Stringent quality (FFPE)
- Fragmentation

Computational burden

- Large file per Gb (especially for raw data)
- Unstandardized algorithms to handle LRS data
- Powerful computing resources required

9° WORKSHOP
IN EMATOLOGIA TRASLAZIONALE
DELLA SOCIETÀ ITALIANA DI EMATOLOGIA SPERIMENTALE
Bologna, Aula "G. Prodi", 19-20 maggio 2025

Sequenziamento long-read: Principi e Applicazioni

Niccolò Bartalucci