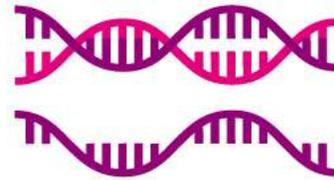




CLINICAL APPLICATION OF NOVEL TECHNOLOGIES



NGS as a diagnostic tool in human diseases

Talk Plot

Next Generation Sequencing – CV and future

Flavours of NGS

Applications in Medicine

- 1. Oncology**
- 2. Reproductive Medicine**
- 3. Pathogen Genomics**
- 4. Hematology**
- 5. AI and NGS**
- 6. Rheumatology**

History of Next Gen Sequencing

1st Generation



Sanger Sequencing

- Infer nt identity with dNTPs and visualize with capillary electrophoresis
- 500-1000 bp fragments
- Expensive, laborious, time consuming

Next Generation (2nd Generation)



Massive Parallel Sequencing

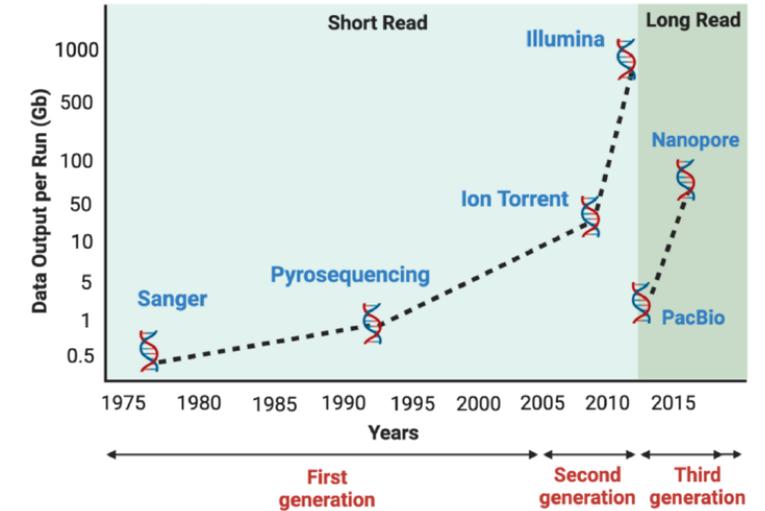
- High throughput from the parallelization of millions of sequencing reactions
- ~50-500 bp fragments

3rd Generation

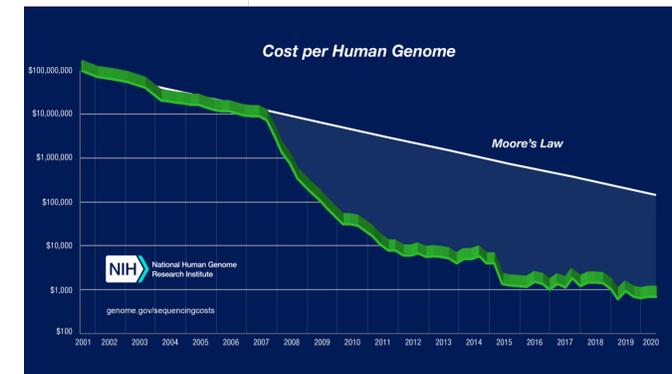


Single Molecule Real Time
Long Reads Sequencing

- Sequence native DNA and/or RNA in real time with single-molecule resolution
- Tens of kb fragments



- Whole Genome Sequencing: • Duration: <48 hrs
• Cost: < 2000 euros
- Whole Exome Sequencing: • Cost: < 300 euros



Next Generation Sequencing Pipeline

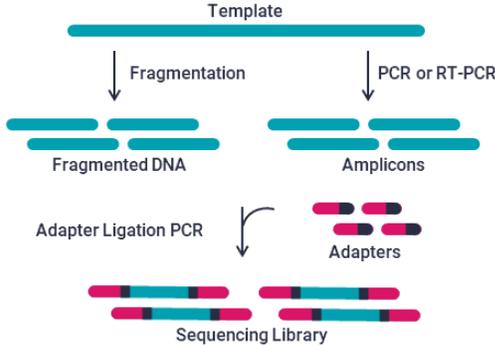
STEP 1: Extraction



STEP 3: Sequencing



STEP 2: Library Prep

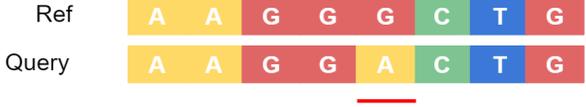


STEP 4: Analysis



Sequence Variants

SNV (Single Nucleotide Variant)

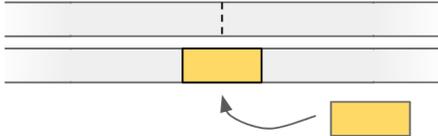


INDEL (Insertion or Deletion)

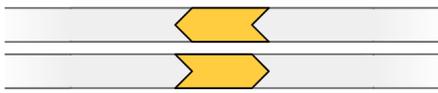


Structural Variants

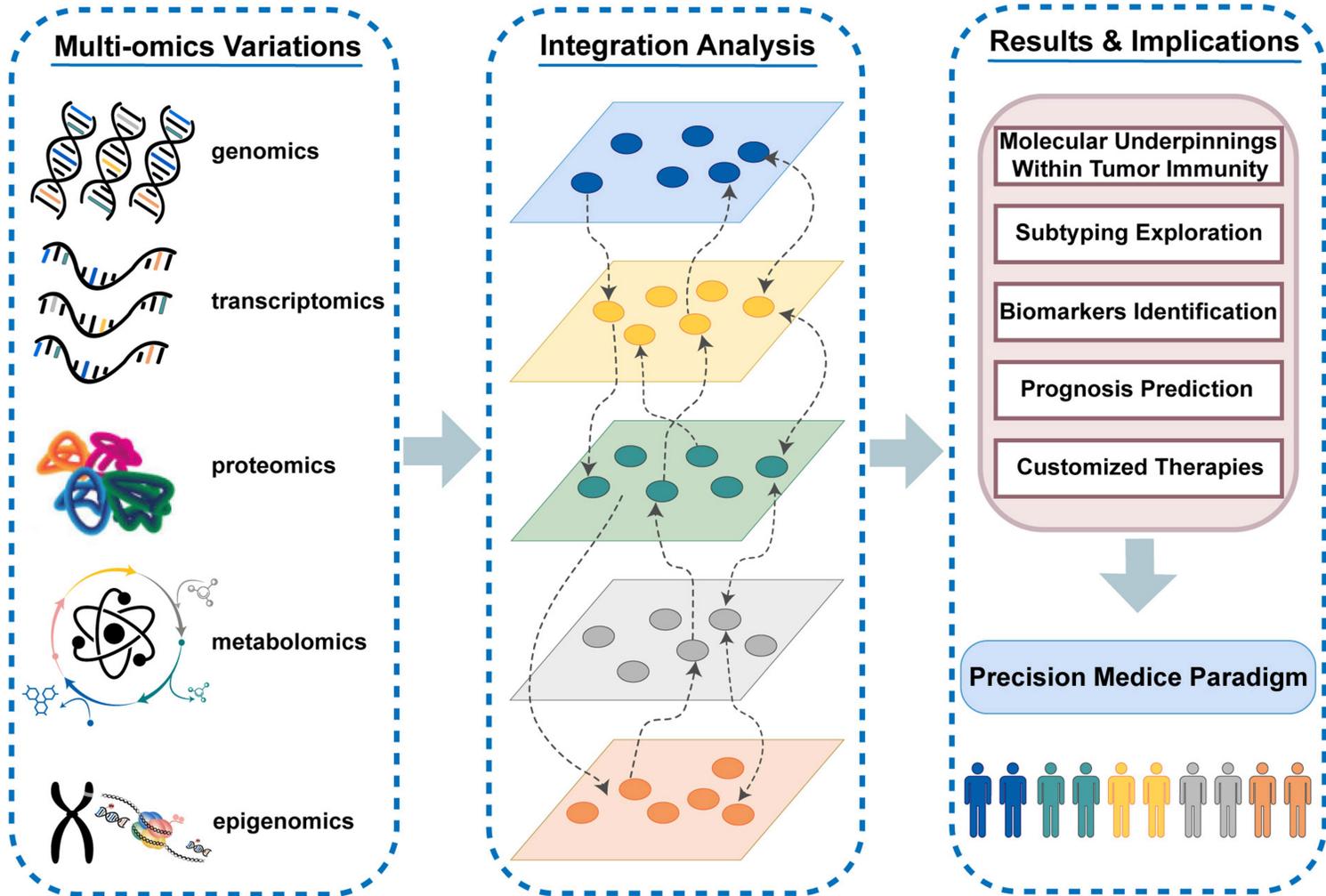
Insertion



Inversion



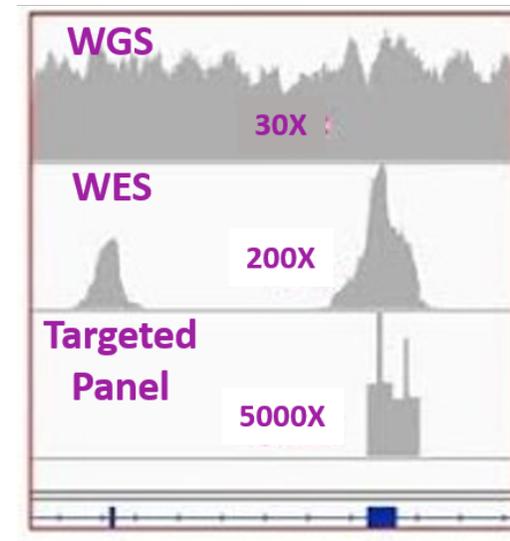
Multi – omics towards personalized medicine



- Different layers of –omics describe the pathophysiology in a multi-parametric manner
- Need for advanced integration!
- Molecular Subtypes of the disease
- Precision Medicine = Personalized Medicine
- Dream or Sisyphean Struggle?

Flavours of genome sequencing

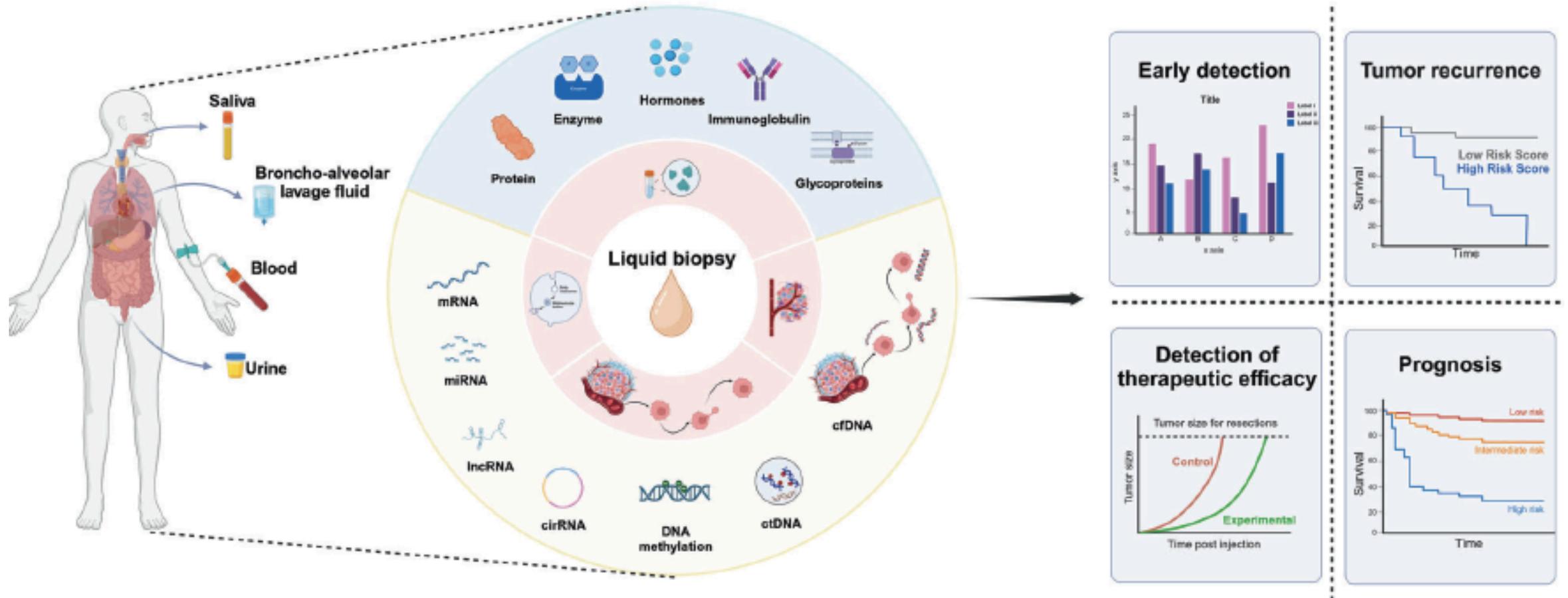
- **Whole Genome Sequencing**: every single bit, depending on depth, de novo assembly
- **Whole Exome Sequencing**: focus on protein-important events, cheaper, weak at regulatory events
- **Targeted Sequencing**: Evolution towards translational research, sequencing for less curious people



- 2% of genome
- 85% of disease causing variants

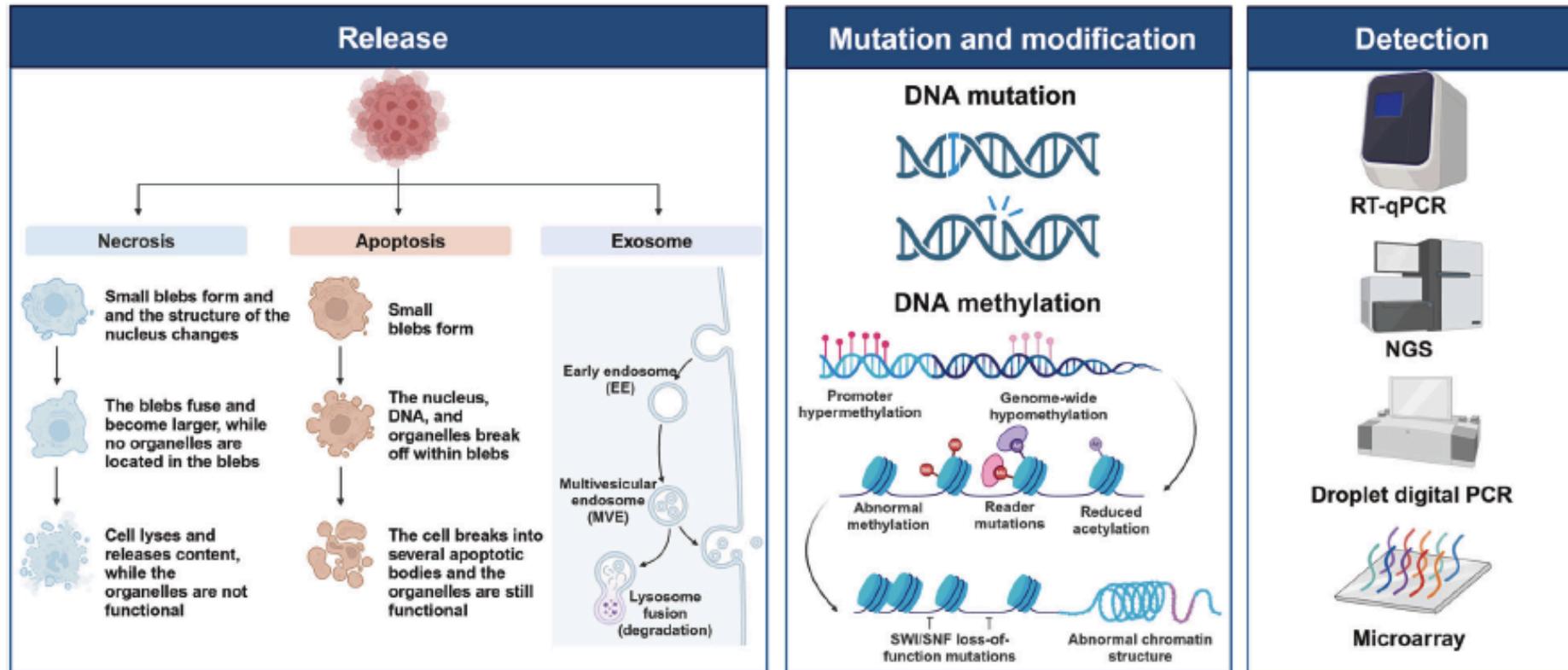
- Only targets of interest
- Seek specific disease causing variants

NGS in liquid biopsies



- Non invasive technique (vs tissue biopsy)
- Markers for early detection, efficacy, refractoriness, prognosis

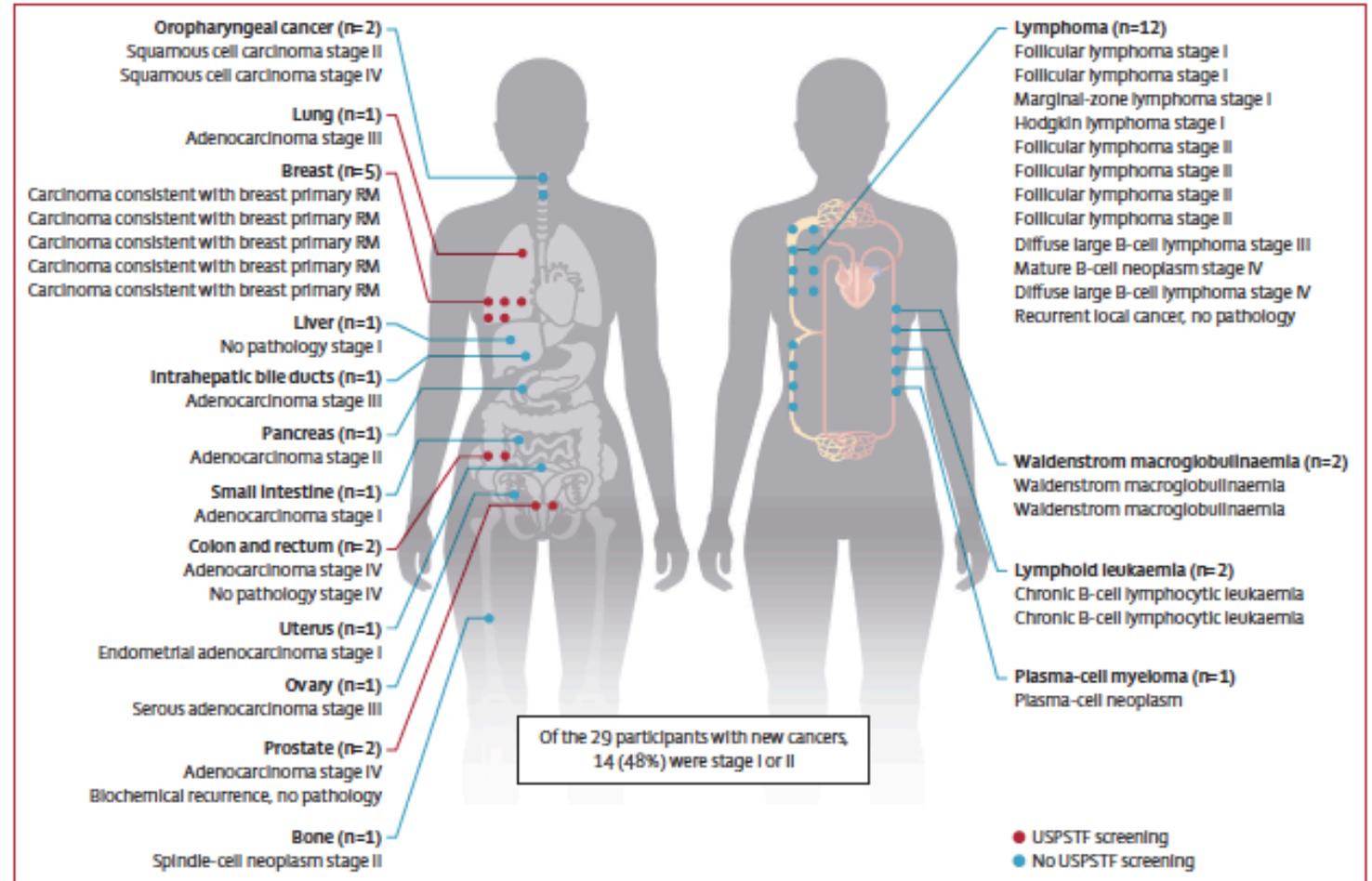
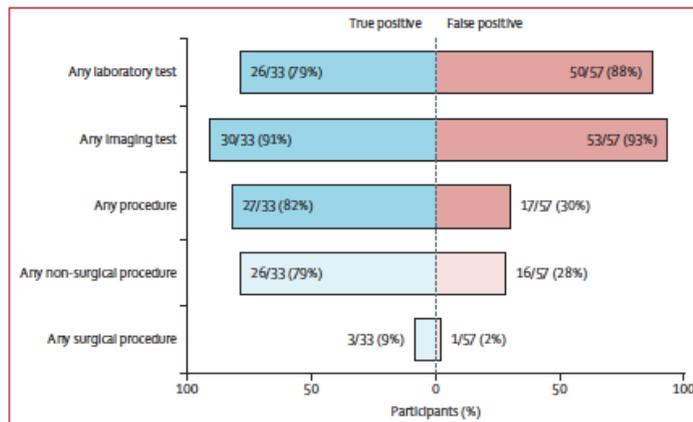
Cancer prognosis and detection through Liquid Biopsy



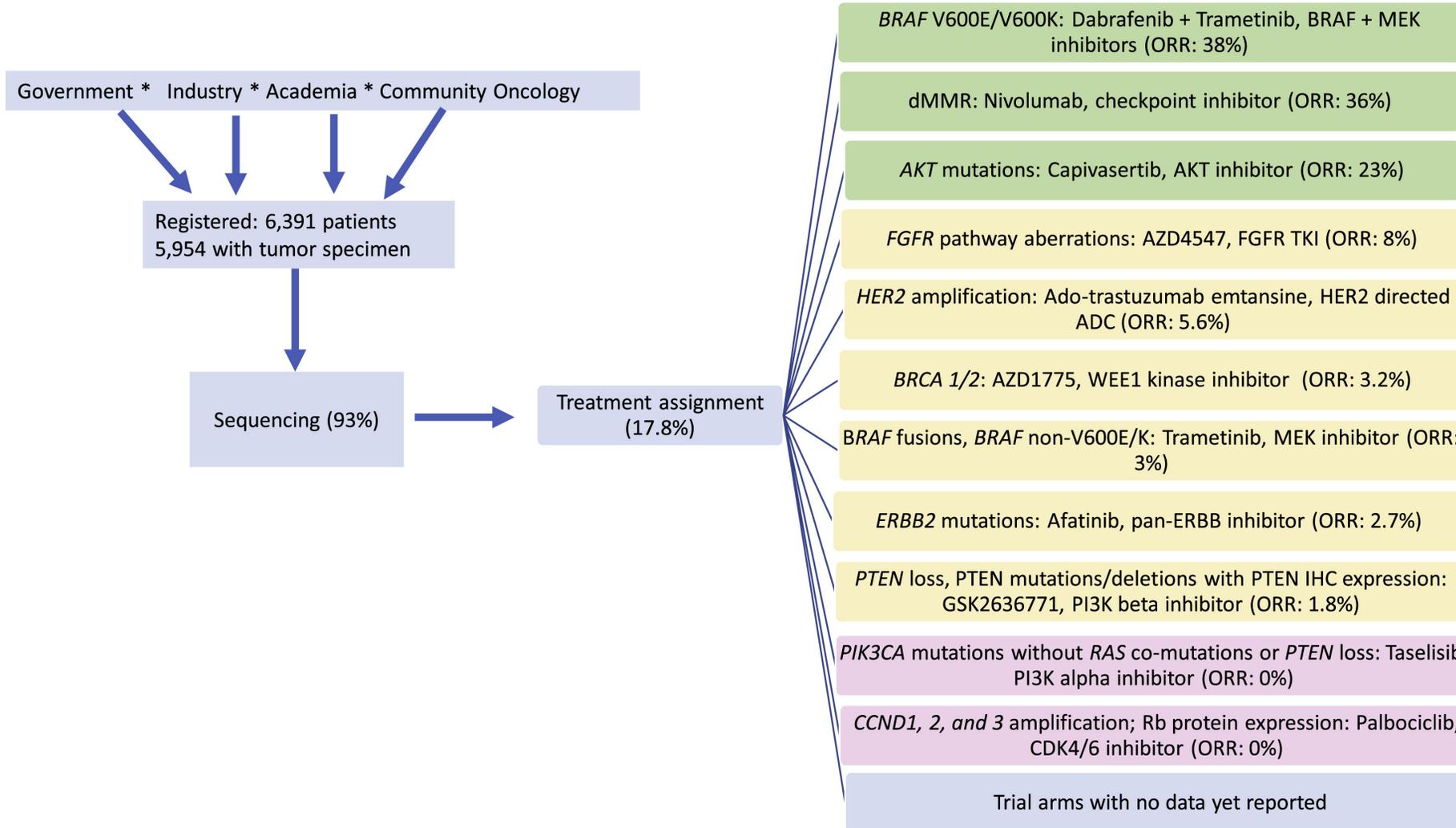
- ctDNA: secreted by tumor cells or released into the circulatory system during the apoptosis or necrosis.
- Mutations and methylation of ctDNA used as detection indicators

Multicancer early detection (MCED) from circulating cell-free DNA (cfDNA) → PATHFINDER

- Prospective cohort study to adults aged > 50 years without signs or symptoms of cancer → MCED testing (methylation pattern with NGS)
- Primary outcome:** time to, and extent of, diagnostic testing required to confirm the presence or absence of cancer
- A cancer signal was detected in 92 where 35 participants were diagnosed with cancer (TPs) and 57 had no cancer diagnosis (FPs)



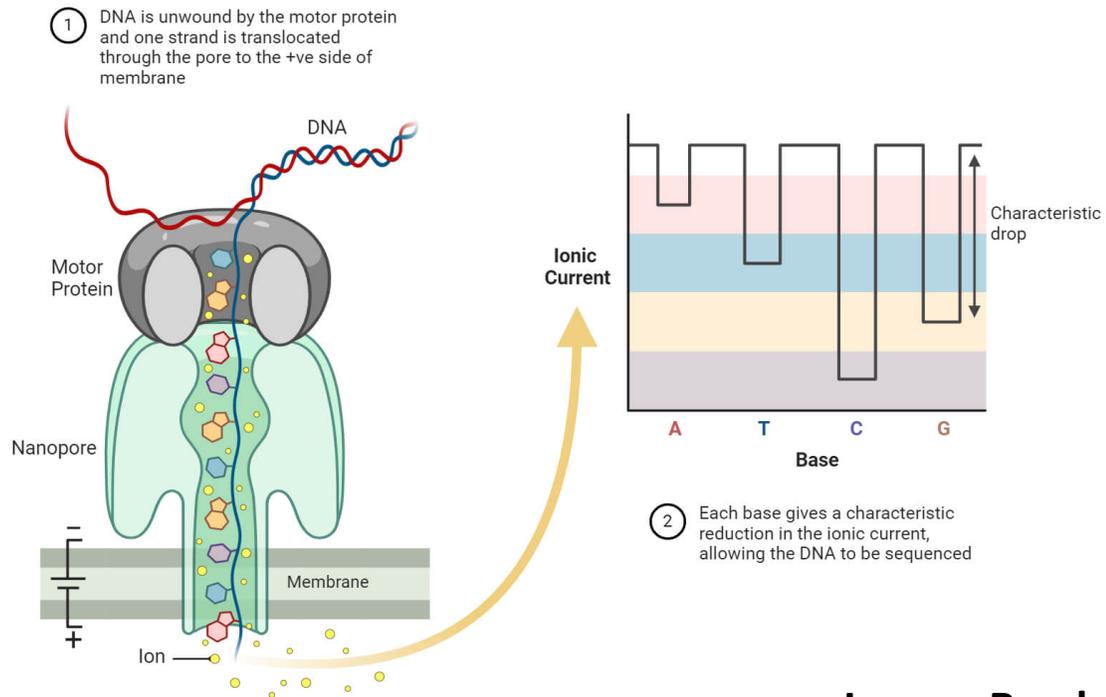
Precision Oncology



- NCI-MATCH (**Molecular Analysis for Therapy Choice**) trial
- Genomically driven, signal-seeking precision medicine
- Treatment-refractory malignant solid tumors
- 6,000 screened with a total of 1,593 patients assigned to 38 substudies.

3rd Generation Sequencing – Long Reads

Nanopore Sequencing Principle

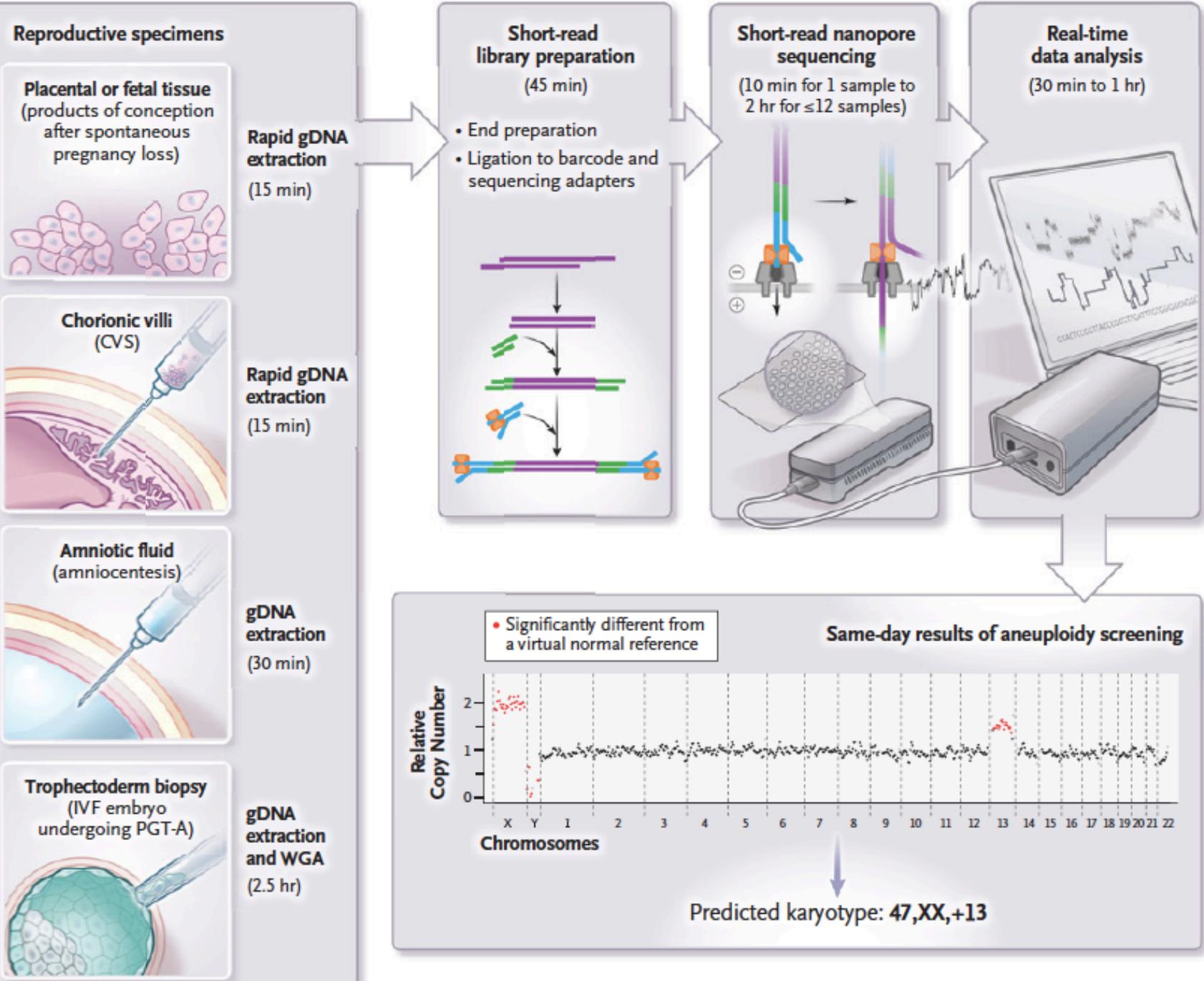


Longer Reads – 50 bp to 4 MB

Structural abnormalities (indels) and CNVs

Prone to errors but real FULL GENOME coverage

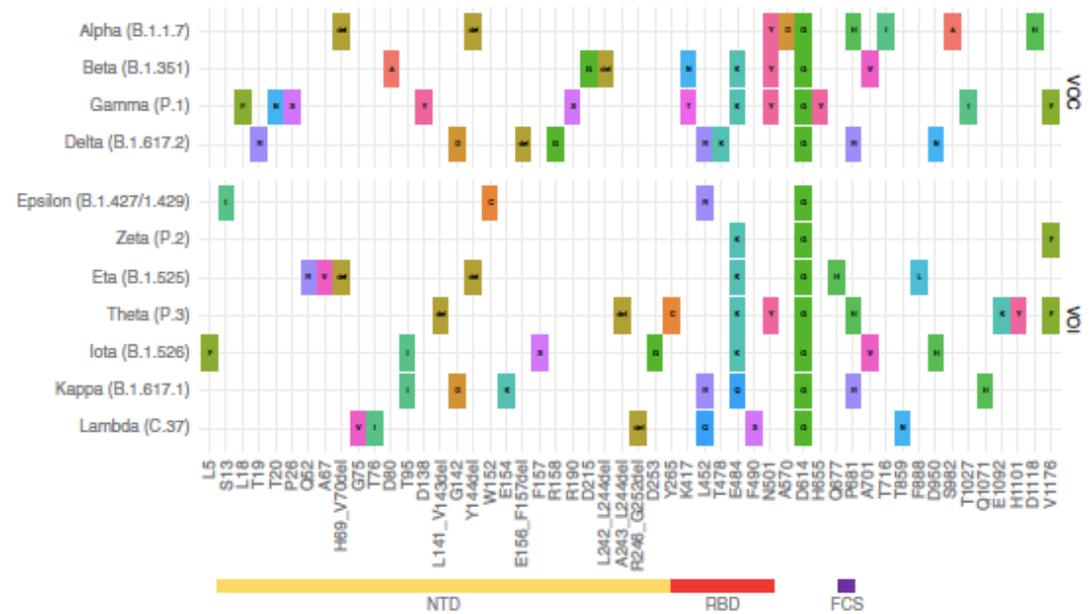
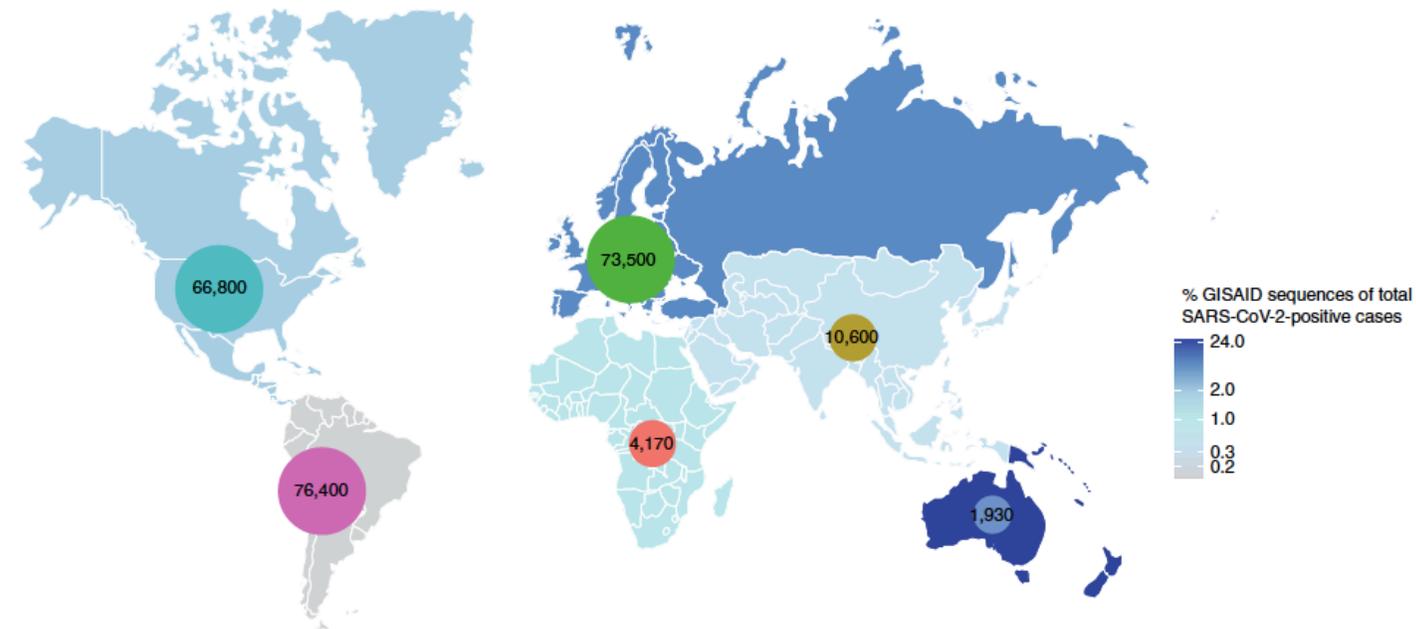
Short-read transpore rapid karyotyping (STORK)



218 sequential, remnant, reproductive specimens of conception after spontaneous pregnancy loss

- Reduced cost,
- Same-day turnaround time
- Perform on-site
- Test for aneuploidy across all chromosomes

Real – time molecular epidemiology



Molecular Surveillance = Genomic Monitoring

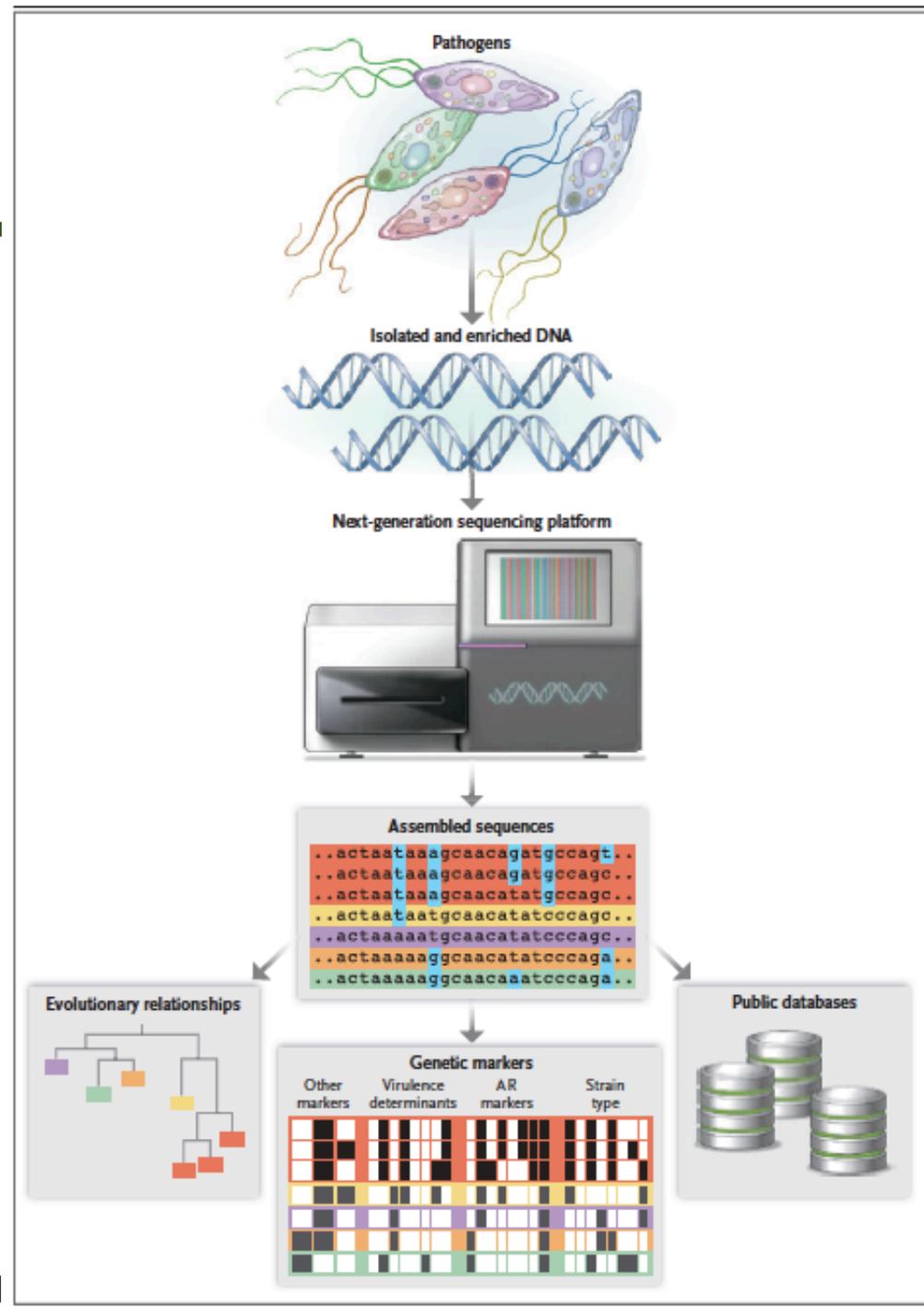
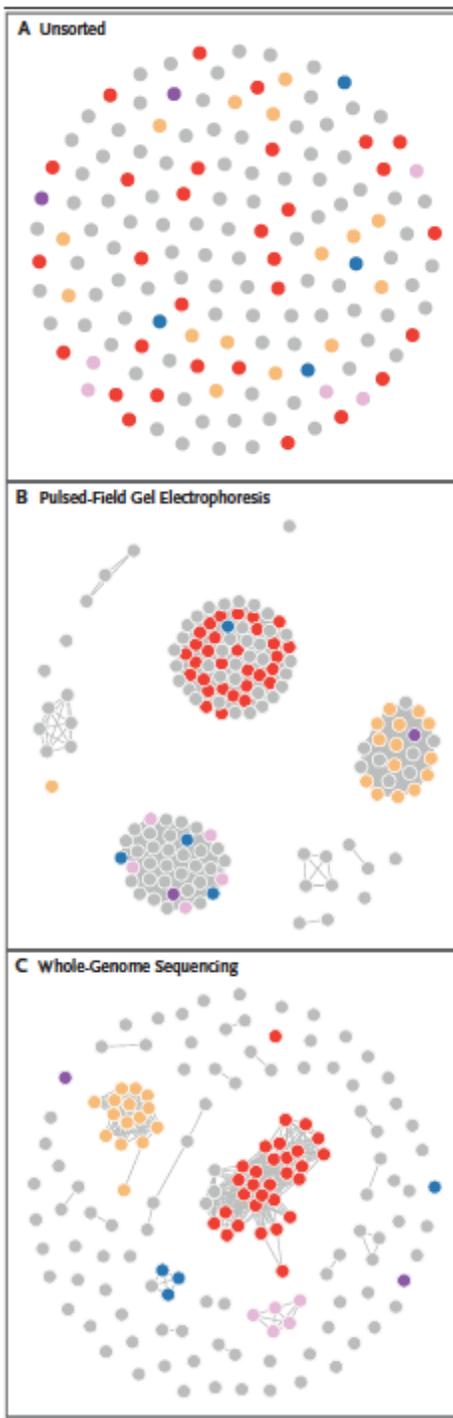
Escapes of therapy – Transmittability

Pathogen Genomics in Public Health

- **Bacterial enteric illness:** improves detection of and response to outbreaks
- **Tuberculosis:** allows better targeting of interventions to stop transmission
- **Legionella:** provides a new tool to understand the ecology of the pathogen in water systems
- **Potential agents of bioterrorism:** allows for improved forensics

**Culture-based system vs NGS from specimens
METAGENOMICS**

Armstrong GL et al 2019 NEJM

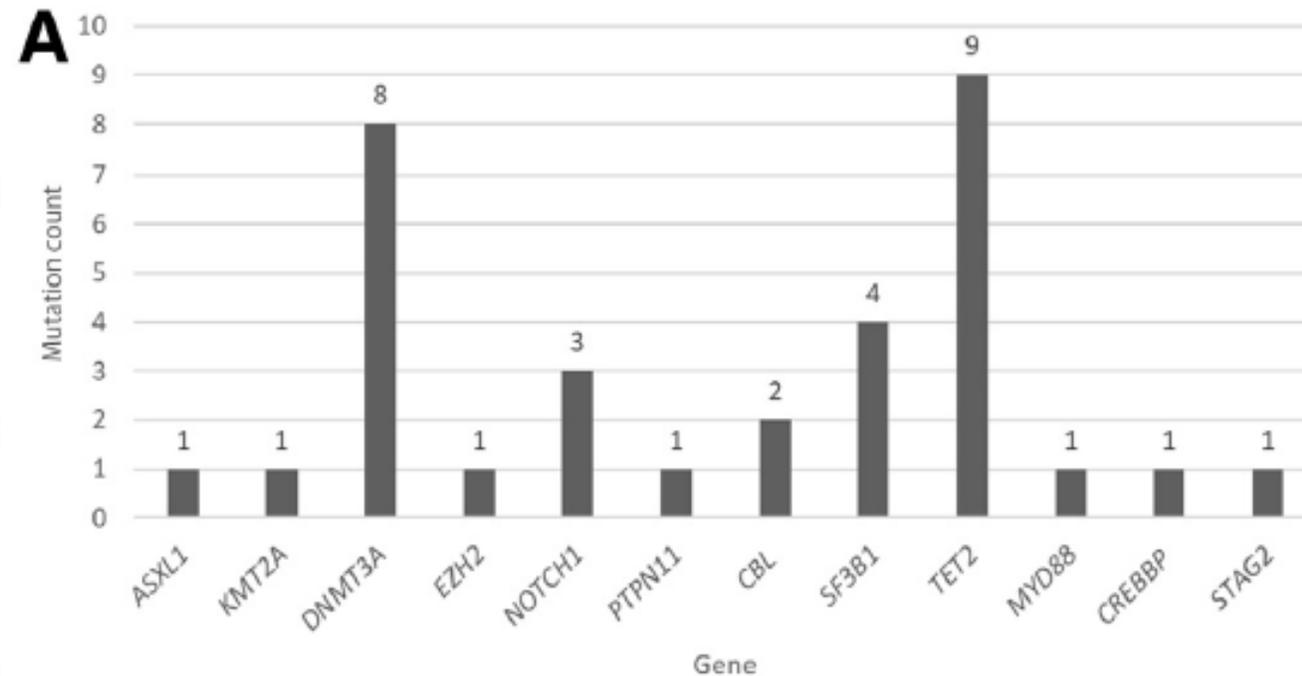
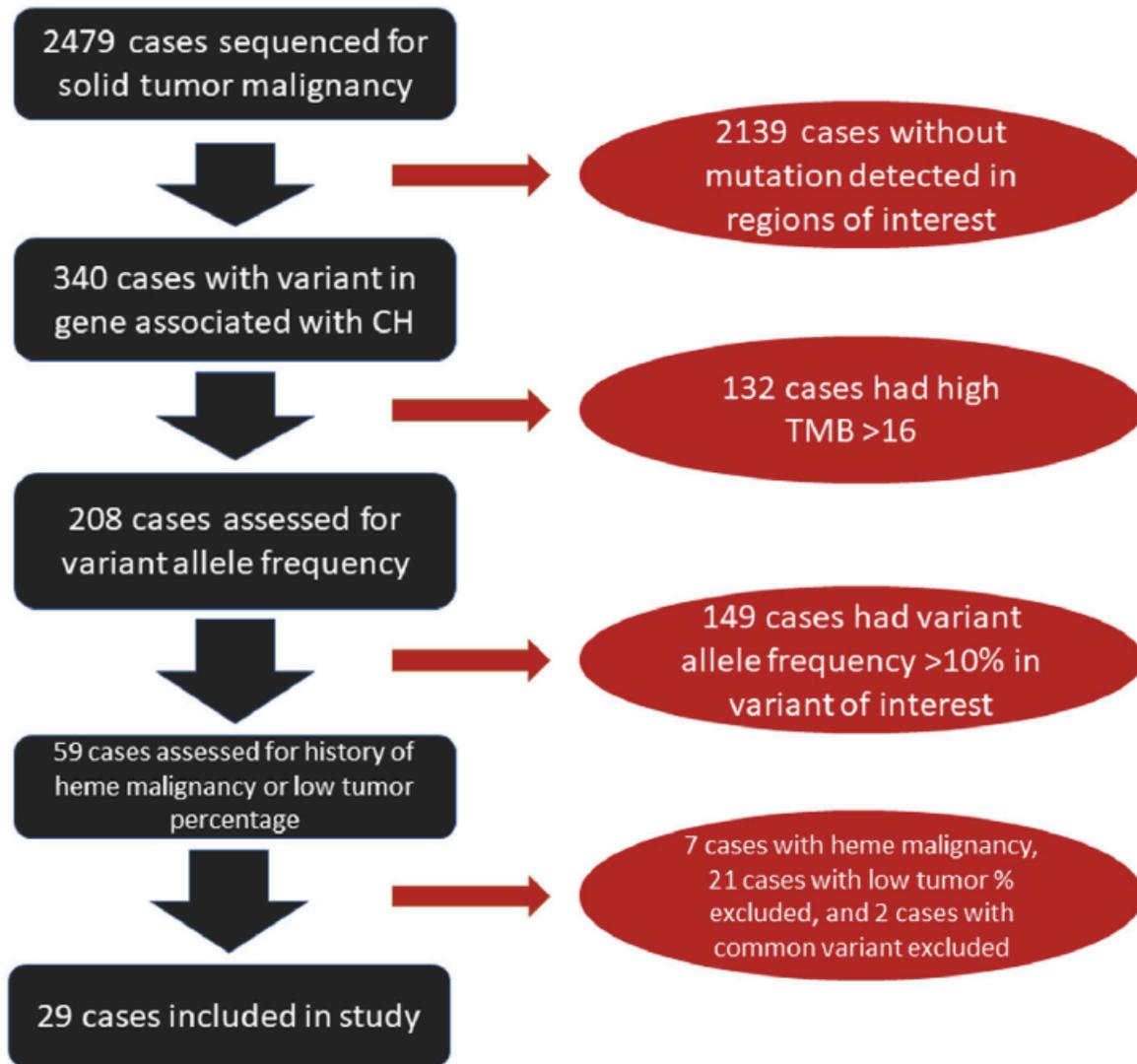


NGS in Hematologic Malignancies

Table 1 Genetic variants detectable by next-generation sequencing assay and of clinical utility in hematologic malignancies

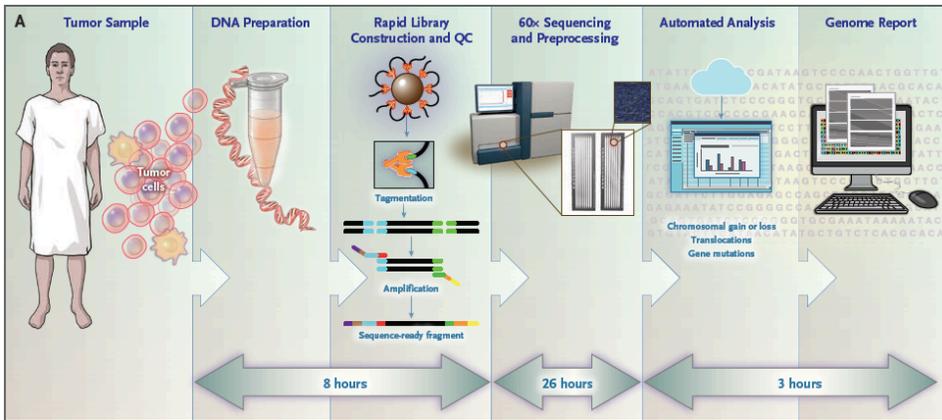
Diseases	Mutated genes	Fusions
Myeloid malignancies	<i>ABL1, ANKRD26, ASXL1^{a,b}, BCOR^{a,b}, BCORL1^a, BRCC3^a, CALR^c, CBL^a, CEBPA^c, CTCF^a, DDX41^c, DNMT3A^a, ETNK1, ETV6^c, EZH2^{b,c}, FLT3^{c,d}, GNAS^a, GNB1^a, IDH1^{a,c}, IDH2^{a,c}, JAK2^{a,c}, KIT^c, KMT2A^d, KRAS^{a,c}, MPL^c, NPM1^c, NRAS^{a,c}, PPM1D^a, PTPN11^{a,c}, RAD21^c, RUNX1^c, SETBP1^a, SF3B1^{a,b,c}, SH2B3, SRSF2^{a,b,c}, STAG2^{b,c}, TET2^a, TP53^{a,c}, U2AF1^{a,b,c}, WT1^c, ZBTB33^a, ZRSR2^b</i>	<i>BCR::ABL1, CFBF::MYH11, DEK::NUP214, KMT2Ar, MECOMr, NUP98r, PML::RARA, RBM15::MRTFA, RUNX1::RUNX1T1,</i>
Lymphoid malignancies ^e	<i>BRAF, CXCR4, CYLD, DIS3, EGR1, FAM46C, FGFR3, HIST1H1E, ID3, IRF4, KRAS, LTB, MAX, MYD88, NRAS, PAX5, RB1, STAT3, STAT5B, TCF3, TP53, TRAF3</i>	<i>ABL1^f, ABL2^f, BCR::ABL1, CRLF2^f, CSF1R^f, DGKHF^f, DUX4r, EPOR^f, ETV6::RUNX1, IGH::IL3, other IGHr, IL2RBr^f, JAK2^f, KMT2Ar, MEF2Dr, MYCr, NTRK3^f, NUTM1r, PAX5r, PGDFRBr^f, PTK2Br^f, TCF3::PBX1, TCF3::HLF, TSLPr^f, TYK2^f, ZNF384r</i>

Clonal Hematopoiesis (CHIP) Variants during Routine Solid Tumor Next-Generation Sequencing

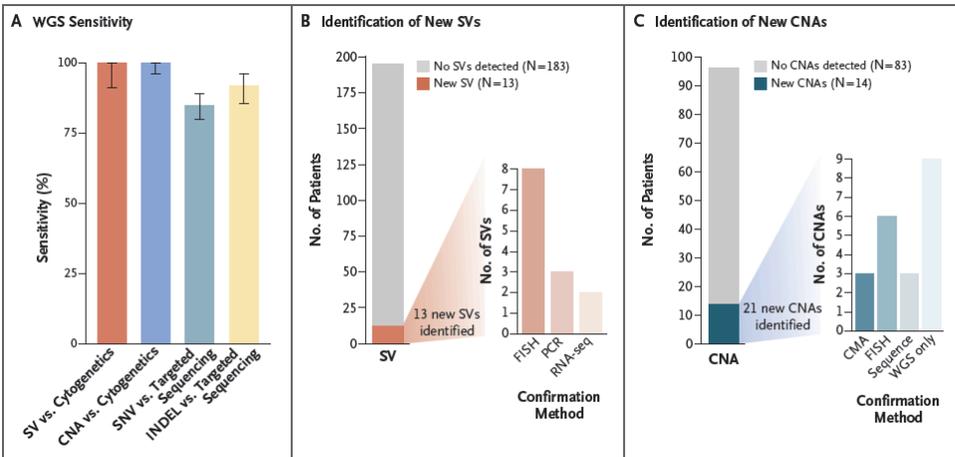
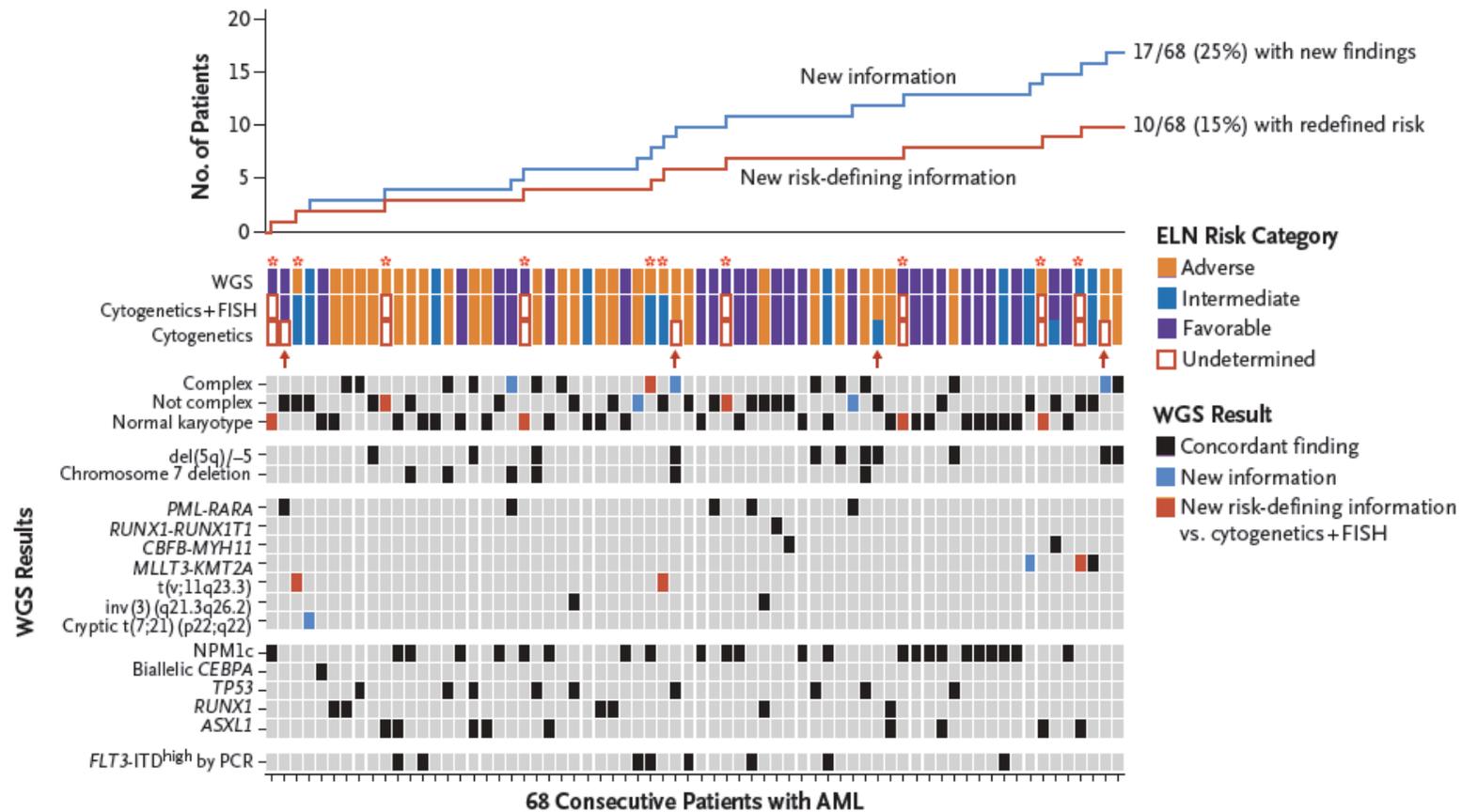


Clonal hematopoiesis (CH) and clonal cytopenia of undetermined significance (CCUS) show independent risk factors for cardiovascular disease and myeloid malignancy

NGS for diagnosis of myeloid cancers

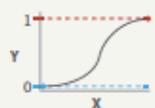


B Diagnostic Yield in 68 Consecutive Patients with AML

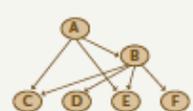


Whole-genome sequencing → rapid and accurate genomic profiling in patients with AML or MDS
 Greater **diagnostic yield** than conventional cytogenetic analysis and more **efficient risk stratification**

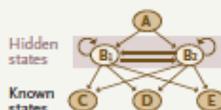
Traditional models



Logistic regression
Probabilistic technique that has been widely used in many models in which the outcome variable is discrete



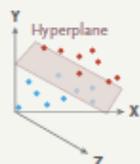
Bayesian networks
Probabilistic graphical models in which a set of variables and their conditional dependencies are modeled with the use of graphs that are directed and acyclic



Hidden Markov models
Probabilistic graphical models that are undirected and may be cyclic



Random forest models
Models that operate through the application of decision trees to iterative subsets of data

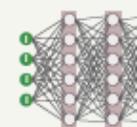


Support-vector machines
Models that separate subgroups according to a hyperplane (separation) in n-dimensional space

Deep-learning models



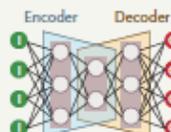
Artificial neural network
Learning models inspired by the organization of the human brain and that comprise nodes called neurons



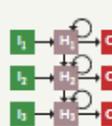
Deep neural network
Neural network with multiple hidden layers of alternating weights and nonlinear activation functions that transform data into useful representations (e.g., for prediction or classification)



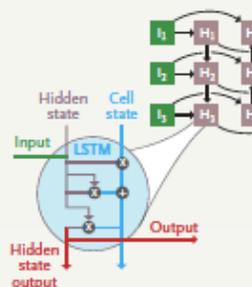
Convolutional neural network
Deep neural network that applies sliding filters over the input data, often an image, to extract relevant features at different spatial locations



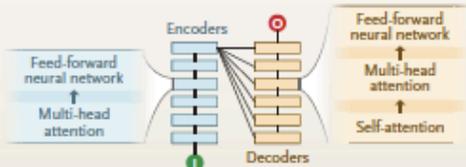
Auto-encoder network
Deep neural network that learns to encode and reconstruct input data to capture essential features



Recurrent neural network
Deep neural network that processes sequential data by using feedback connections that allow information to persist over time during the modeling process



Bidirectional recurrent neural network with LSTM units
Complex recurrent neural network that effectively models sequential data by processing it in both forward and backward directions, enabling it to capture information from distant elements in the sequence through LSTM units

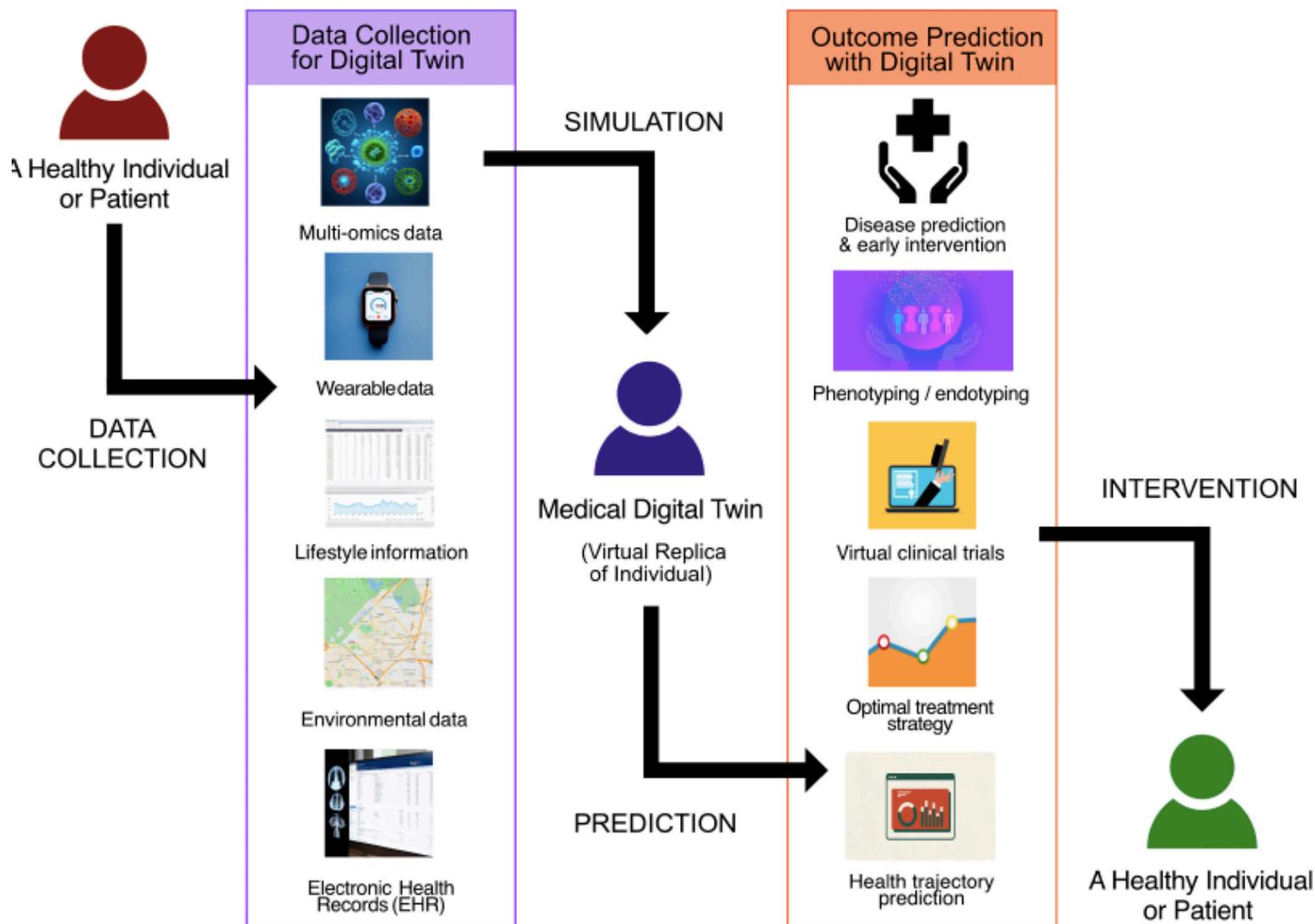


Transformer models
Deep neural networks that use self-attention mechanisms to capture relationships between different elements in a sequence

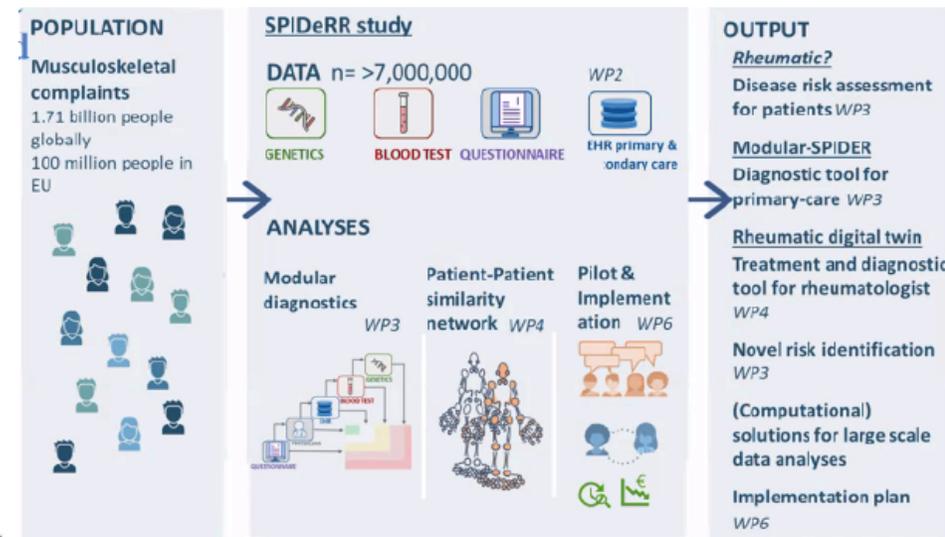
AI in Molecular Medicine

- Multi-omics data are complex before the description of pathophysiology
- Analyze down to abstract and low dimensionality data
- Traditional and Deep Learning Methods
- Hidden motifs
- Accuracy in variant calling
- Prioritization of variants for rare diseases – Causal for Mendelians
- Haplotype for clinical management

The idea of the Digital Twin

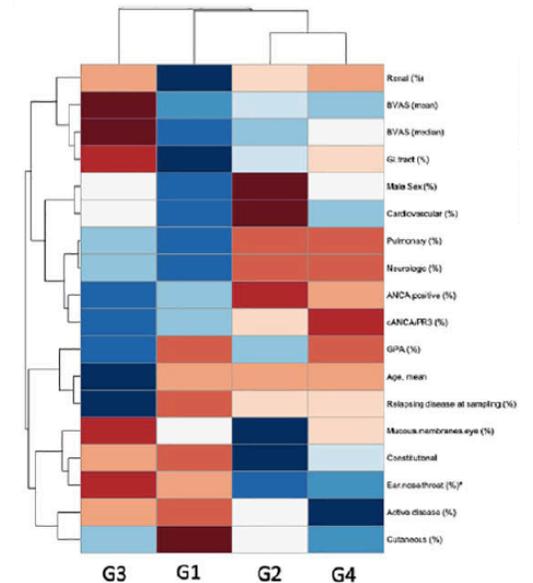
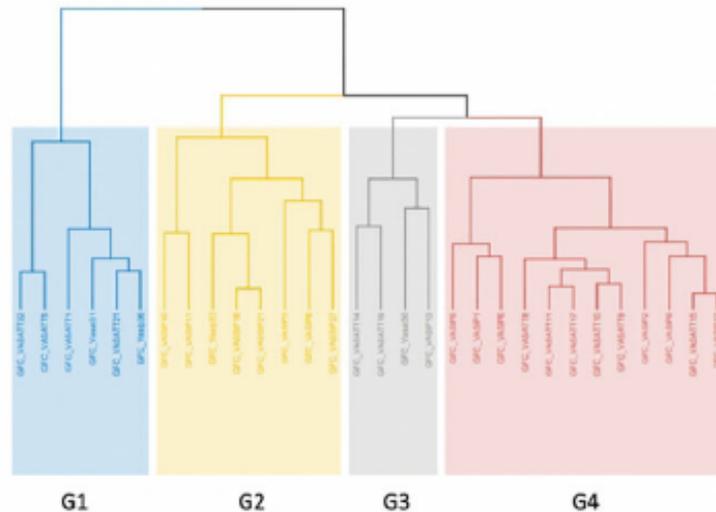
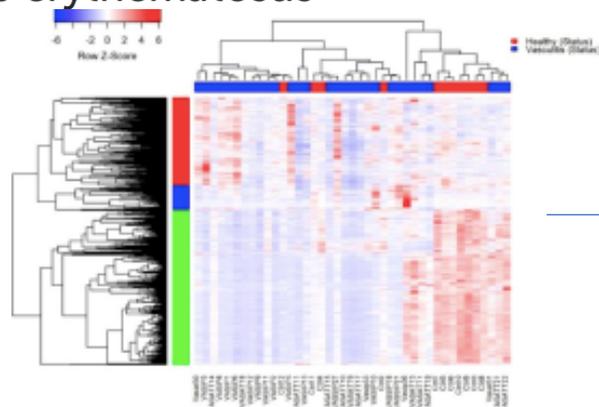


- AI training datasets concluding with next gen data
- **Predictive models to twin every new individual to a phenotype universe**
- Paradigm → Musculo-skeletal complaints



Whole Blood Transcriptome Profile establishes Molecular Endotypes of the Disease

The genomic landscape of ANCA-associated vasculitis: Distinct transcriptional signatures, molecular endotypes and comparison with systemic lupus erythematosus



Unsupervised molecular taxonomy analysis → **4 endotypes** with neutrophil degranulation, aberrant metabolism and B-cell responses as potential mechanistic drivers.

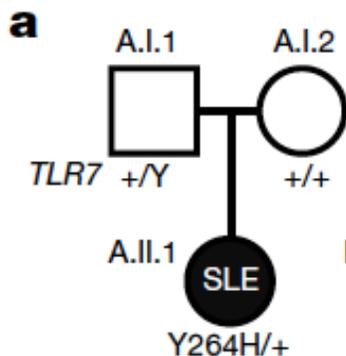
Therapeutic insights directed towards B-cells, complement cascade, type I IFN or drug repurposing

Genetic variations of TLR7 drives SLE phenotype

Article

***TLR7* gain-of-function genetic variation causes human lupus**

- TLR7 → sensor of degraded genetic material
- **Case** → Female from Spain with SLE diagnosis since 7th year of age
- Refractory ITP, high titer ANA/anti-dsDNA, low complement, arthralgias, choreia
- Past Therapies → RTX, GCs, AZA, MMF, ETN



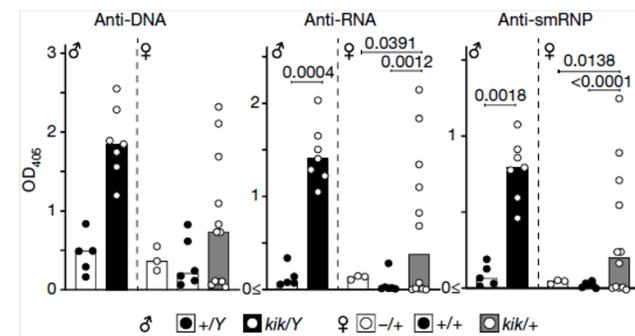
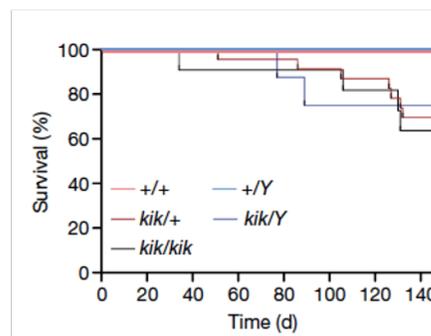
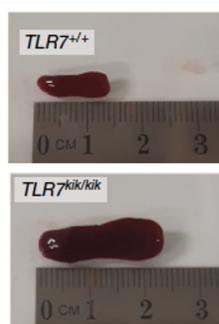
Whole exome sequencing → De novo mutation

TLR7 p.Tyr264His (Y264H) missense variant (GOF)

Physicians do not want to hear about mice

Insertion of mutated allele in mouse model with CrispR/Cas9

Autoimmunity α



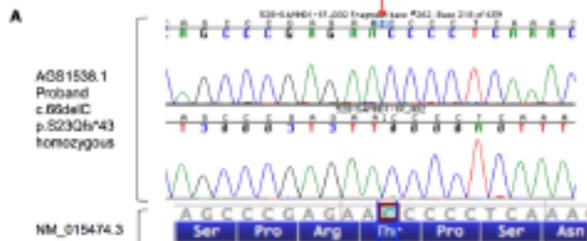
Increased GC B cells, ABC cells, pDCs, Tfh cells

Whole exome sequencing in rare pediatric cases

Increased interferon I signaling, DNA damage response and evidence of T-cell exhaustion in a patient with combined interferonopathy (Aicardi-Goutières Syndrome, AGS) and cohesinopathy (Cornelia de Lange Syndrome, CdLS)

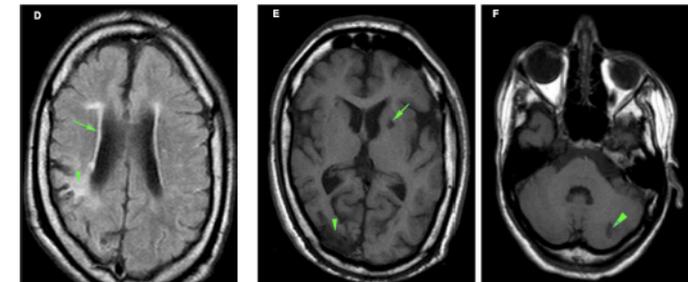


Sorina Boiu^{1,2,3,4}, Nikolaos Paschalidis⁵, George Sentis⁶, Theodora Manolakou⁶, Andrianos Nezos⁷, Manolis Gialitakis⁶, Maria Grigoriou⁶, Erato Atsali¹, Melpomeni Giorgi⁸, Argirios Ntinopoulos⁹, Clio Mavragani⁷, Periklis Makrythanasis^{5,10,11}, Dimitrios T. Boumpas^{6,12} and Aggelos Banos⁶



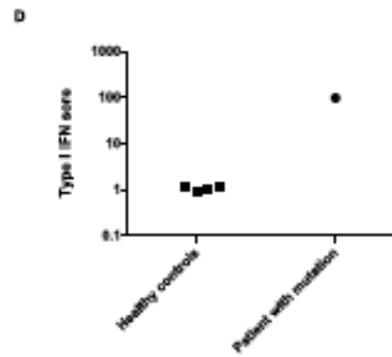
B

Whole-Exome Sequencing	
Reads	190,513,256
Mapped	187,891,382
Exonic Variants	23,626
Splicing Variants	1,304
Synonymous SNV	11,740
Nonsynonymous SNV	10,807
Stopgain	92
Stoploss	10
Frameshift deletion	114
Frameshift insertion	93
Nonframeshift deletion	195
Nonframeshift insertion	118



C

Chr	Start	End	Fuse	Gene	ExonicFuse	AAChange	Obs	Ref	Zyg	AD
chr20	35579680	35579981	exonic	SAMHD1	frameshift deletion	SAMHD1:NM_015474:exon1:c.88_87T	A	AG	hom	0-113
chrX	63489294	63489284	exonic	SMC1A	nonsynonymous SNV	SMC1A:NM_006306:exon22:c.1339C>A;p.M1172K	T	C	hem	1-155



- 20yo man → chilblain lesions, resorption of distal phalanges, somatic and psychomotor retardation, microcephaly, synophrys, hearing loss
- Whole exome sequencing → 2 genetic lesions (AGS/CdLS) exhibits distinctive features of genomic damage and interferon responses

Take Home messages

- NGS is rapidly showing the potential to substitute conventional laboratory diagnostics
- Whole Genome Sequencing yields high accuracy information but still disconnected of clinical practice due to cost/logistics
- Precision Medicine is built on next generation sequencing – Milestone yet to come
- Artificial Intelligence is connecting the missing parts for big –omics data
- Therapeutic resolution for rare diseases



Thank you!