Translating Scientific Breakthroughs from the Genomics Revolution

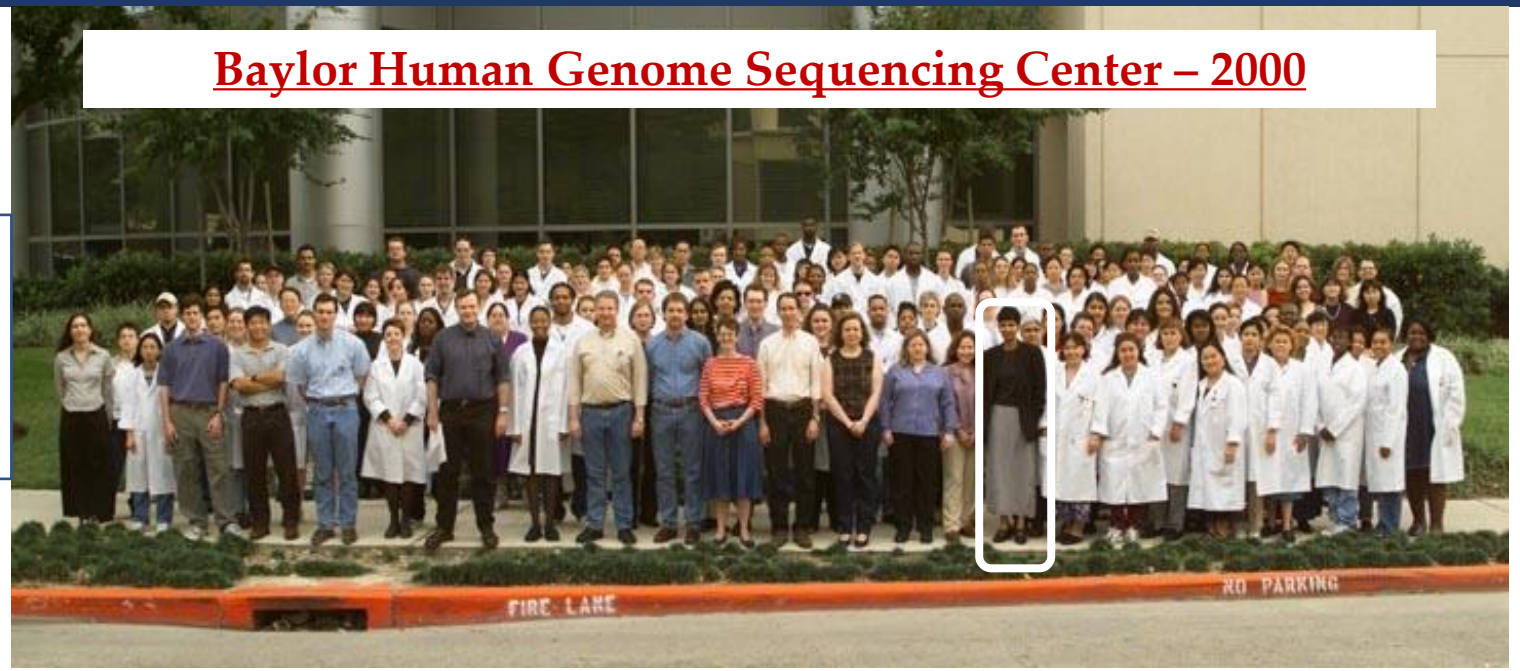
Preethi Gunaratne, Ph.D.
John & Rebecca Moores Professor
Department of Biology & Biochemistry,
Director, University of Houston Sequencing Core
University of Houston

SEASR 2024 Annual Meeting
Nashville, TN

The University of Houston Sequencing and Gene Editing Core

"Student Run, Comprehensive, Sample to Data Insight Pipelines for Commercial and Academic User Groups"

- Human Chromosomes 3, 12, X
- 15,000 FL-cDNA Sequences Mammalian Gene Collection



Baylor Human Genome Sequencing Center – 2000

1999

- ~1000 people
- ~300 Instruments

- 9 Months
- $1B
- 1 Human Genome

2020

UHSEQ Sequencing Core

Illumina NextSeq 2000

- Effort of 1 Person
- 1 Instrument

- 48 hrs.
- $1000/Genome
- 4 Human Genomes

2024

Illumina NovaSeq X+

- Effort of 1 Person
- 1 Instrument

- 24 hrs.
- $200/Genome
- 4 Human Genomes

Illumina NovaSeq X+

UHSEQ Sequencing Core

"Niche Sequencing"

microRNA

1. Developing microRNA biomarkers and therapeutics for cancer
2. Discovering microRNA-based blood biomarkers for predicting individuals at high risk for Opioid Overuse Disorder (OUD)

Non-Canonical Dark Matter

Platform for designing peptide & mRNA vaccines from immunogenic neoantigens from RNA fusions for the treatment & prevention of cancer

16S Sequencing

A Field-Tested Platform for Extracting Personalized DNA Biomarkers for Predicting Drainage Heights, % Contribution, Well-to-Well Communication for Precision Well Placement
Providing Metrics for Asset Development

Comprehensive, Sample to Data Insight Pipelines for Commercial Academic and User Groups

Preethi Gunaratne, Ph.D.
Director

UHSEQ Sequencing Core

Nuwan Acharige, Ph.D.

Lead Scientist
16S Sequencing Platform

Chathurani Ranatunge, Ph.D.

Lead Analyst
Breast Cancer Prevention Vaccine Platform

Sakuni Rankothgedera

Lead Scientist
Single Cell and Spatial Transcriptomics Platform

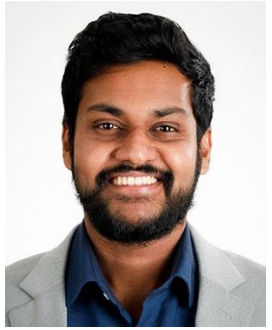
Shiyanth Thevasagayampillai

Lead Scientist
mRNA Vaccine Discovery Platform

Aaranyah Kandasamy

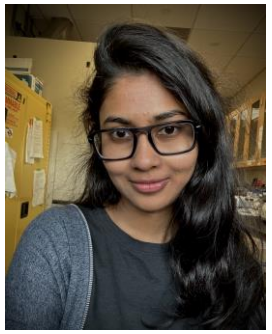
Lead Scientist
PDX -Actionable Mutation Discovery Platform

Dilshan Adikari



NGS Specialist
RNA & DNA Sequencing Platform

Nirmali Samarakoon



NGS Specialist
RNA & DNA Sequencing Platform

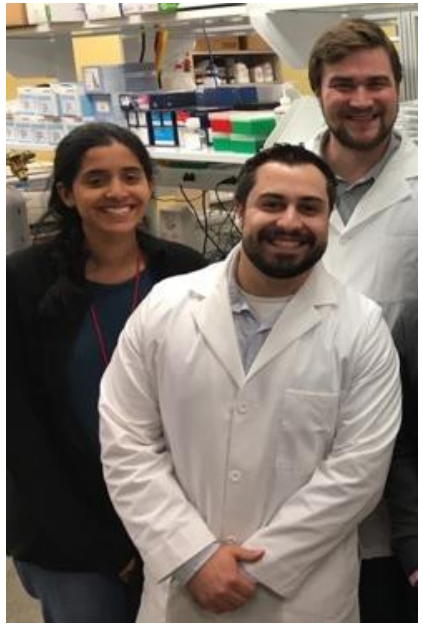
Enes Dasdemir

Single Cell Sequencing Hematologic Cancers

Cole Woody

***In silico* fusion validation platform**

Ashley Benham-Duret, Ph.D.
Senior Scientist, Sales Executive II,
10X GENOMICS



Asha Palat, Ph.D.
Scientist,
Novogene

Ian Wilson, M.Sc.
Scientist,
ThermoFisher SCIENTIFIC

Micah Castillo, Ph.D.
Spatial Field Application Scientist
10X GENOMICS

Kimberly Holloway, Ph.D.
Associate Director,
Translational Oncology,
Iterion Therapeutics

Brandon Mistretta, Ph.D.
Field Application Scientist,
10X Genomics

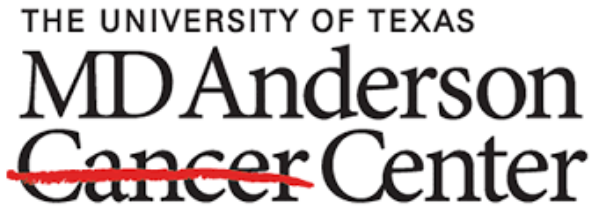
Sujash Chatterjee, PhD
Lead-Product Marketing & Commercialization, Genomics,
QIAGEN

Sujash Chatterjee, PhD
BioPharma Strategy and Commercialization
illumina

Jignesh Chandrana, M.Sc.
Lead Lab Automation Developer
Invitae Genetic Testing

“We provide custom solutions to your problems and perform optimization until we crack the code”

- **Industrial Sequencing**
- **Commercial Sequencing**
- **Health Care**
- **Academic Labs**


THE UNIVERSITY OF TEXAS
MD Anderson
Cancer Center


UTHealth
Houston


Baylor
College of
Medicine


THE TEXAS HEART INSTITUTE


utmb Health


HOUSTON
Methodist
LEADING MEDICINE


UNIVERSITY of
HOUSTON


TEXAS A&M UNIVERSITY
HEALTH


Texas Children's
Hospital


TEXAS TECH
UNIVERSITY


AUGUSTA
UNIVERSITY


LEHIGH
UNIVERSITY


Queen's
UNIVERSITY


UF
UNIVERSITY of
FLORIDA


Cleveland Clinic


NEW YORK UNIVERSITY


COLUMBIA COLLEGE
COLUMBIA UNIVERSITY in the CITY of NEW YORK

Industrial Clients

The University of Houston Sequencing and Gene Editing Core

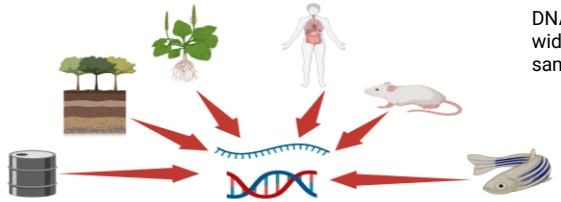
"Student Run, Comprehensive, Sample to Data Insight Pipelines for Commercial and Academic User Groups"

UHSEQ Sequencing Core

Experimental Design and Assay Optimization

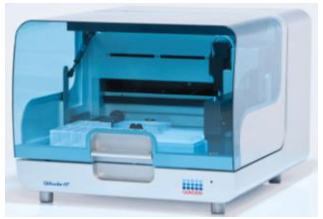
- **Consult with our team prior to each assay to ensure optimal results.**
  - **Streamline your projects to avoid future repeats**

Sample Preparation and Nuclei Acid Extraction



DNA/RNA Extraction from a wide variety of fresh/preserved sample types.

QIACube: 12 samples/hour
Compatibility: DNA/RNA



QIACube HT: 96 samples/hour
Compatibility: DNA/RNA

Sample Quality Assessment

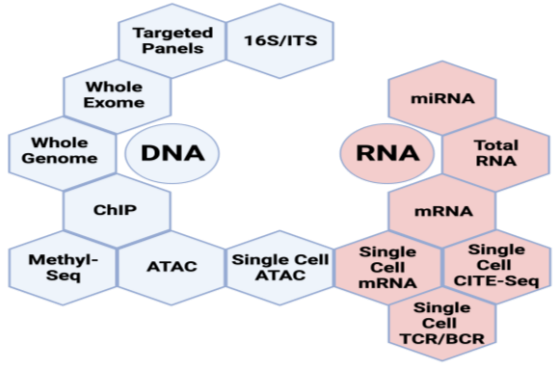
Qubit Flex Fluorometer
Throughput: 8 samples/run, 96 samples/10 minutes
Compatibility: DNA/RNA Protein
Application: Sample Concentration

Agilent 4200 TapeStation System
Throughput: 96 samples/1.5hours
Compatibility: DNA/RNA
Application: Sample Quality, Size distribution, RINe, DV200

Multi-Assay RT-PCR

Agilent Aria MX Real Time PCR System
Throughput: 96 samples
Compatibility: SYBR/ROX/FAM
Application: Sample Concentration, Gene Expression Detection, Mouse/Human Tissue Contamination Detection (PDX)

Next Generation Sequencing



Illumina MiSeq
Maximum Read Output: 25M
Maximum Cycle Output: 600

Illumina NextSeq 2000
Maximum Read Output: 1.1B
Maximum Cycle Output: 600

Single Cell Sequencing

10X Genomics Instrumentation

- **A) Chromium Controller:** 3'/5' Gene Expression, Immune Profiling, ATAC, CITE-Seq
- **B) Chromium X:** Fixed RNA, High-throughput, All other assays.
- **C) Visium CytAssist:** Spatial profiling of RNA and Protein on FFPE or Fresh-Frozen tissue slides.
- **D) Chromium Connect:** Full automation of 10X pipeline.

Industrial Collaborations

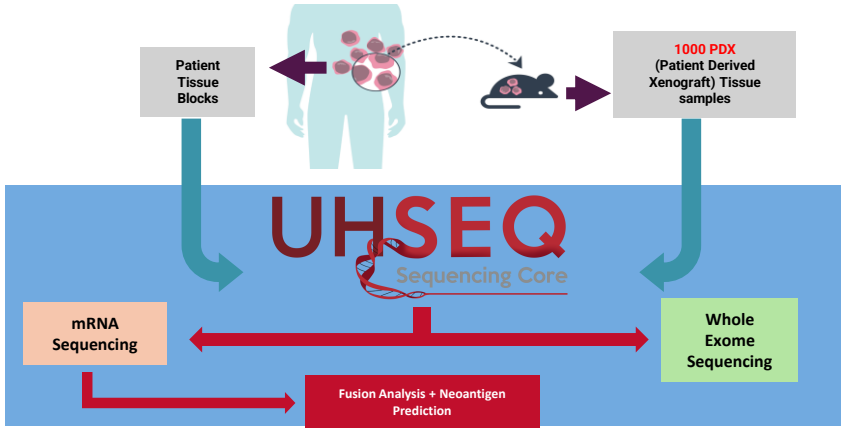
Cancer

xenSTART

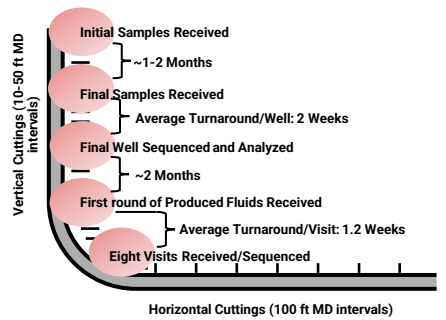
Energy

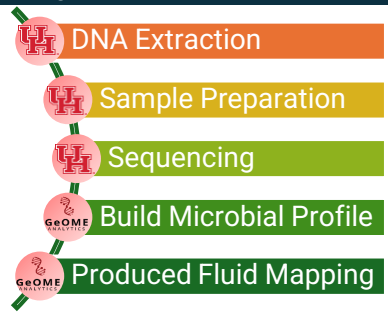
GeOME ANALYTICS · Chevron · Shell

Actionable Mutations in PDX Models



Reservoir Drainage Diagnostics





Auxiliary Instrumentation

Countess 3 FL
Compatibility: RFP + GFP
Application: Cell Counting

Transblot Turbo Transfer System
Application: Western Blotting

Funding

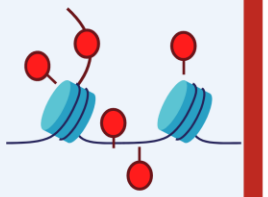
- **Large Core Equipment Grant – 2021**
- **Large Core Equipment Grant Program -2022**

Genomics

- Whole Genome Seq
- Whole Exome Seq
- 16S rRNA Seq
- ITS Seq
- Metagenomics

Epigenomics

- ChIP-seq
- ATAC-seq
- Hi-C Assay
- Methyl Seq



Transcriptomics

- mRNA-seq
- Total RNA-seq
- miRNA-seq

Single-cell Sequencing

- Single-Cell Gene Expression (3',5' GEX)
- Single-Cell Immune Profiling (5' VDJ)
- Single-Cell ATAC
- Single-Cel Multiome ATAC + GEX
- Single-Cell Gene Expression Flex

Bioinformatics Analysis

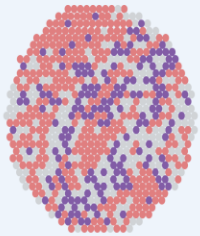
Analysis available for all of the services provided. Our team can create custom analyses tailored to your project goals upon request

UHSEQ Sequencing Core

SERVICES WE PROVIDE

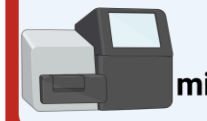
Spatial Transcriptomics

- Visium CytAssist
- Visium HD

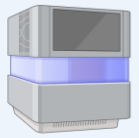


Sequencing Only

Sequence pre-made libraries with us!



miSeq



Nextseq 2000

Bulk RNA Seq Vs Single Cell RNA Seq

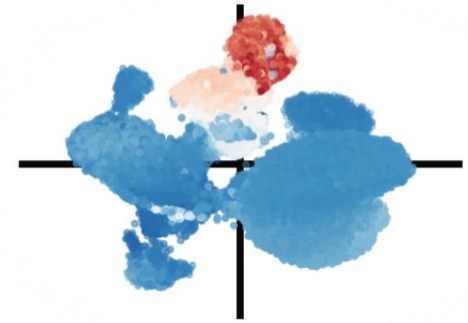
10x GENOMICS

Single-Cell
Suspension & Processing

Single-Cell
Expression Profiles

Gene A
Gene B
Gene C
Gene D
...

Cell-Type Identification



Single Cell Gene Expression

A.

Patient Biopsy

- tumor site
- adjacent tissue
- lymph node
- peripheral blood

B.

Bulk-RNA
Processing

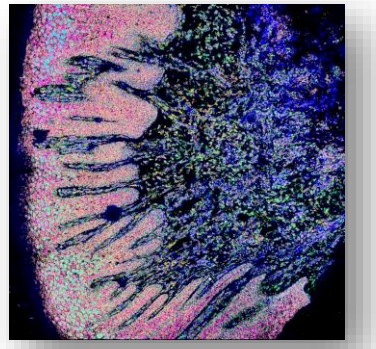
Averaged Expression
of Sample

Gene A
Gene B
Gene C
Gene D
...

?

Tissue Section

10x Visium Spatial Transcriptomics



Spatially Resolved Gene Expression

Adapted from 10x Genomics

Discovery of Immunogenic Neopeptides from Actionable Fusions

to Develop Peptide & mRNA Vaccines for Cancer Treatment &

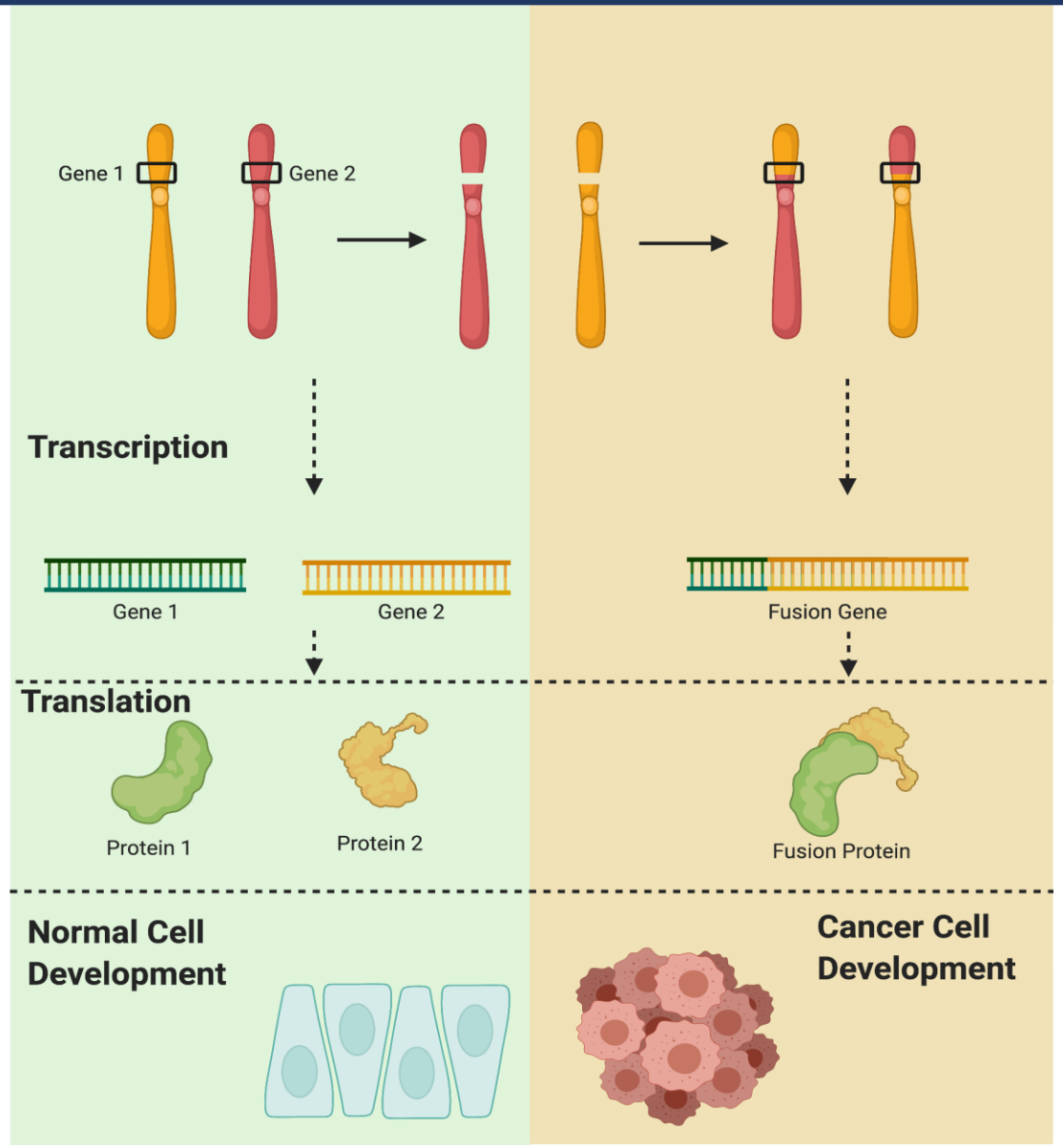
Micah Castillo, Ph.D.

UNIVERSITY of HOUSTON

College of Natural Sciences and Mathematics

UHSEQ Sequencing Core

Gene Fusions



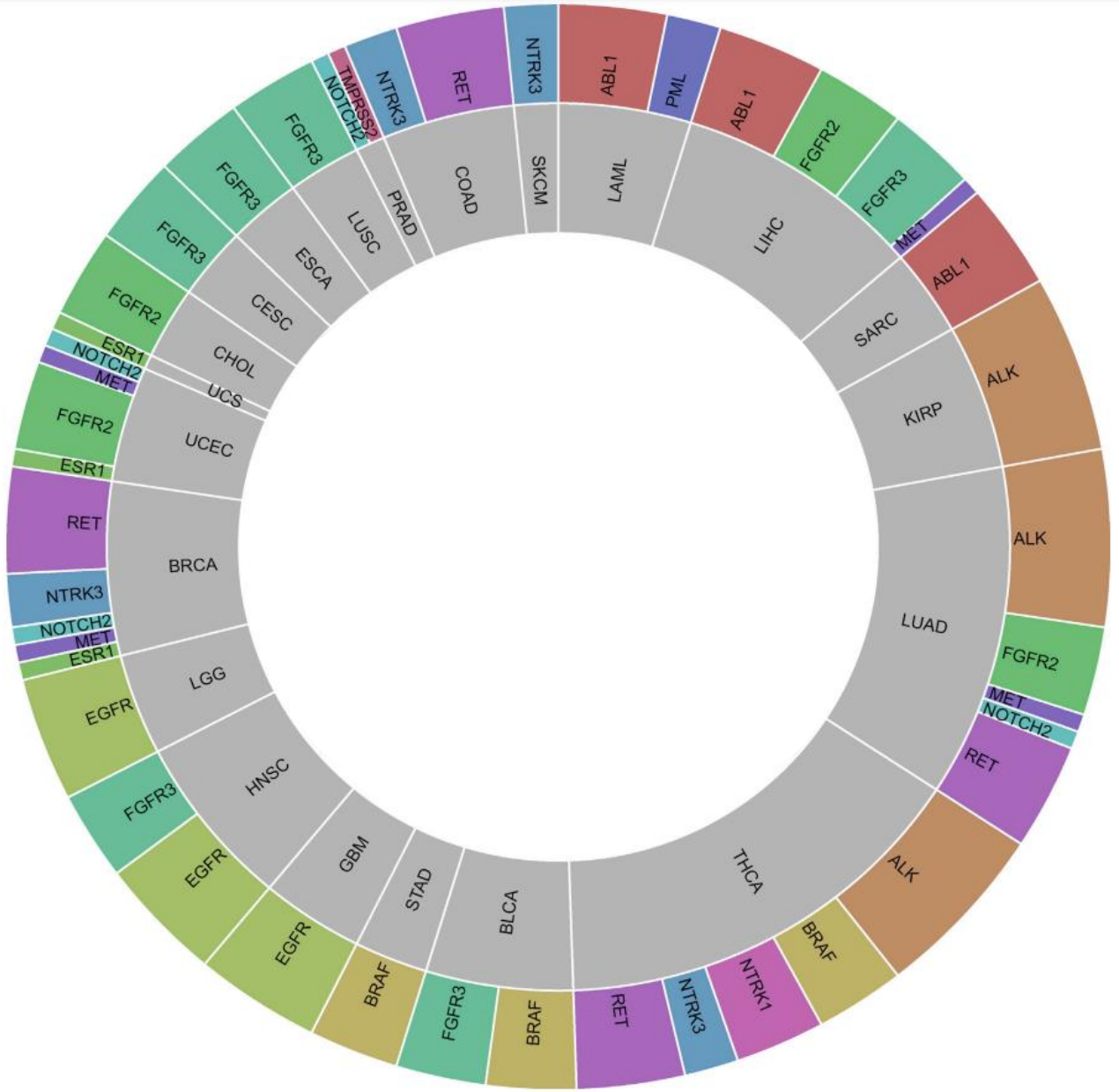
Created with Bio Render

Importance in Cancer

- Structural chromosome rearrangements result in the exchange of coding or regulatory DNA sequences between genes.
- Fusions may also occur on the RNA level as a result of aberrant splicing.
- Many fusions are strong driver mutations in neoplasia.
- **BCR-ABL fusion gene is a major marker in chronic myelogenous leukemia (CML).**
- Identifying fusions may lead to development of novel therapeutics or early detection methods.

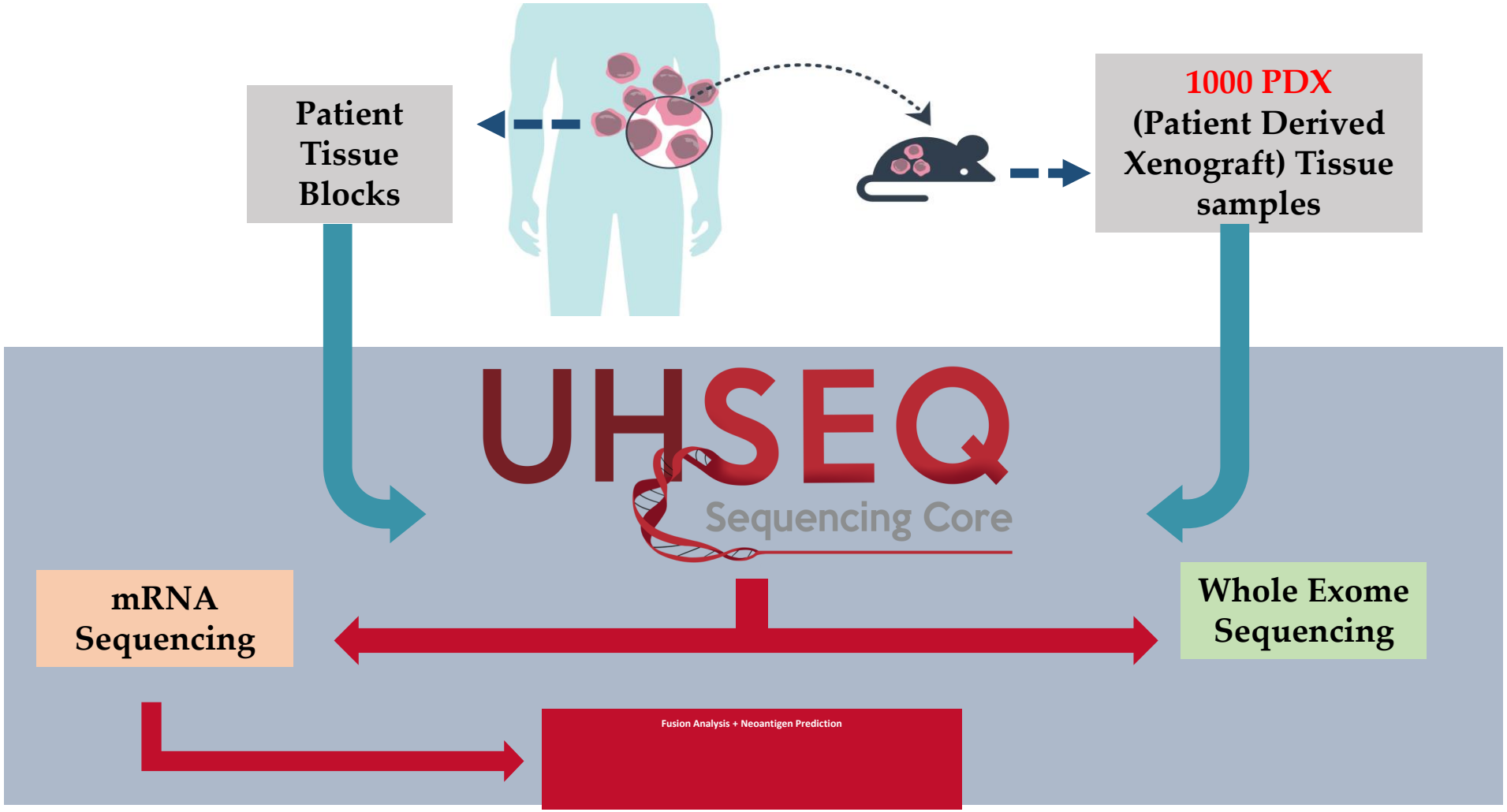
TCGA Actionable Fusions List

Gene	Drug(s)	Fusion Partner(s)
ABL1	Bosutinib, Dasatinib, Imatinib, Nilotinib, Ponatinib,Venetoclax	FUBP3, BCR
ALK	Alectinib, Brigatinib, Ceritinib,Crizotinib,Entrictinib, Lorlatinib, MEK inhibitors, novel ALK inhibitors, PF2341066 , TAE684	EML4, STRN
BRAF	Cobimetinib, MEK inhibitors , Selumetinib, Sorafenib, Trametinib	MRPS33, SND1
EGFR	Afatinib , EGFR TKIs , Erlotinib,first-generation and irreversible EGFR TKIs , gefitinib , HSP90 inhibitors , ZD6474	SEPT14
ESR1	Anti-estrogens	CCDC170, TMEM2
FGFR2	AZD4547, BGJ398, Debio1347, Erdafitinib, FGFR inhibitors	ATE1, BICC1, SHTN
FGFR3	AZD4547, BGJ398, Debio1347, Erdafitinib, FGFR inhibitors	TACC3
MET	Crizotinib	CAV1, ST7
NOTCH2	Gamma secretase inhibitors	SEC22B
NTRK1	Crizotinib,Entrectinib, IGF1R inhibitors, Larotrectinib, pan-TRK inhibitor	IRF2BP2, TPM3
NTRK3	Entrectnib, Larotrectinib, Midostaurin	ETV6
PML	Arsenic Trioxide, Tretinoin, Volasertib	RARA
RET	BLU-667, Crizotinib, LOXO-292, Nintedanib, Sunitinib,Vandetanib	CCDC6, ERC1, NCOA
TMPRSS2	DNA-PKc inhibitors	ERG, ETV4



Vellichirammal et al. (2020), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7044684/pdf/main.pdf

Actionable Mutations in PDX Models



Discordant Reads

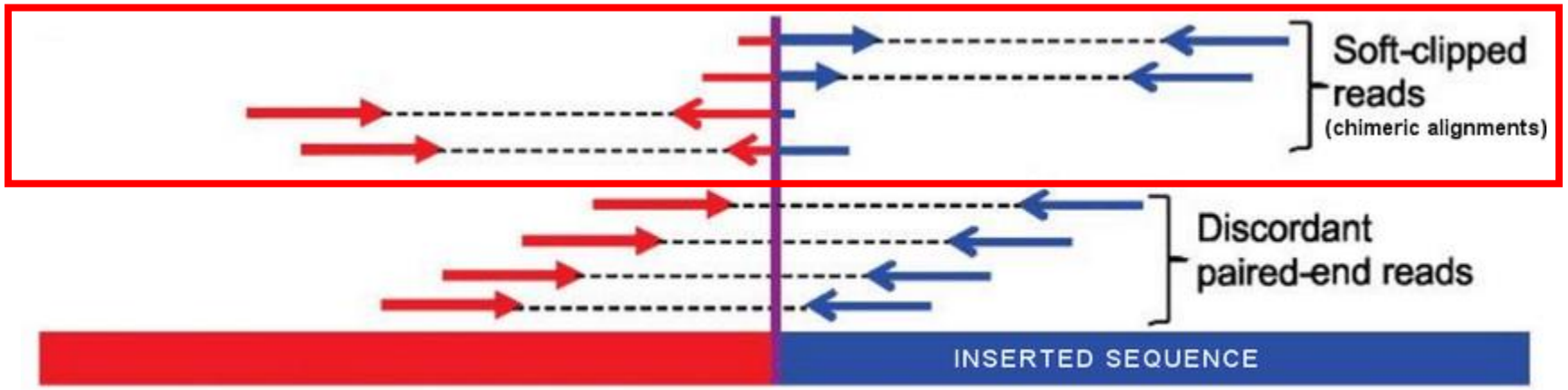
Paired-end reads with unexpected mapping distances.

Fusion callers that use only discordant reads

- **VariantMeta Caller (DNA fusions)**
  - Gezsi et al. VariantMetaCaller: automated fusion of variant calling pipelines for quantitative, precision-based filtering. *BMC Genomics*. 2015
- **ChimeRScope**
  - Li et al. ChimeRScope: a novel alignment-free algorithm for fusion transcript prediction using paired-end RNA-Seq data. *Nucleic Acids Research*. 2017
- **ChimPipe**
  - Rodrigeuz-Martin et al. ChimPipe: Accurate detection of fusion genes and transcription-induced chimeras from RNA-seq data. *BMC Genomics*. 2018

50% False Discovery Rate

- Mismapping of read pairs to different gene family members and pseudogenes
- Mismapping reads containing repeats shared by two different genes



Zapparoli 2017

Chimeric RNA Discovery Pipeline using Multiple Fusion Callers

Library Preparation of NSCLC Patient Samples and Next Generation Sequencing

RNA Seq Data from Lung Cancer Patient Samples

Identification of gene fusions through CLC Genomics Workbench 20, Dragen Fusion Caller, and EasyFuse Fusion Caller

Aligned to Human Reference Genome Hg38	Detect Fusion Gene tool re-mappings the unaligned ends of reads (CLC)	Refine Fusion tool re-counts the number of fusion crossing reads(CLC)

Filter out Fusions Overlapping in Adjacent Normal Samples

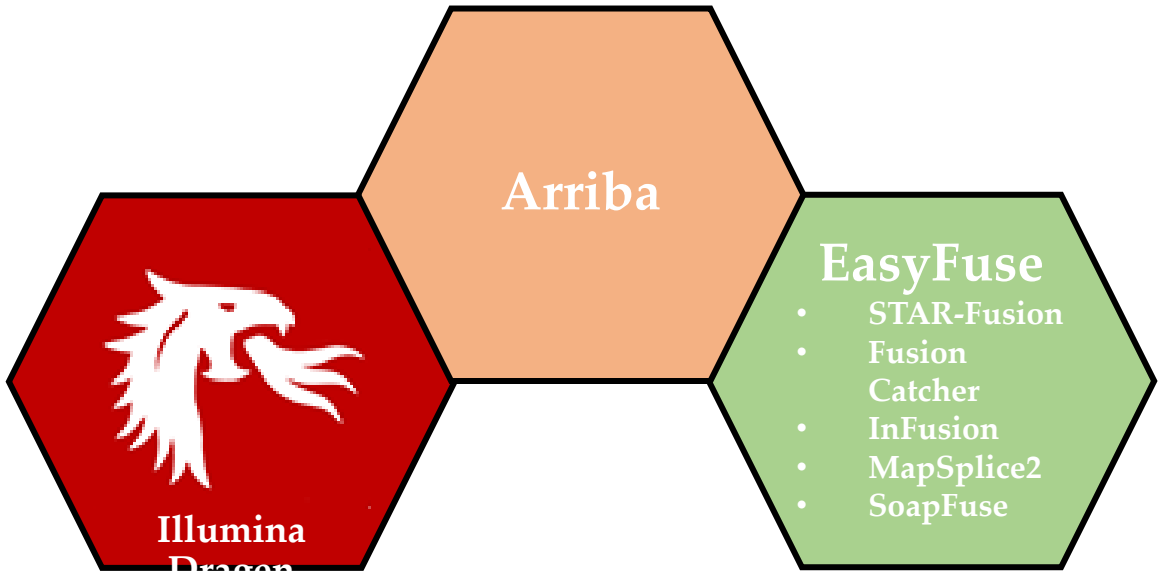
Subtract fusions represented in both normal and tumor patient samples

Overlap with TCGA dataset to assess actionability

Novel Fusions reported by a 1000 sample TCGA Breast Cancer Dataset run by the University of Chicago's fusion caller.

Fusion Discovery Pipeline

Higher Stringency, Higher Specificity



Arriba - Uhrig *et al.*, 2021, (*Genome Research*)
EasyFuse - Weber, *et al.*, 2022, (*Nature Biotechnology*)

Lower Stringency, Higher Sensitivity

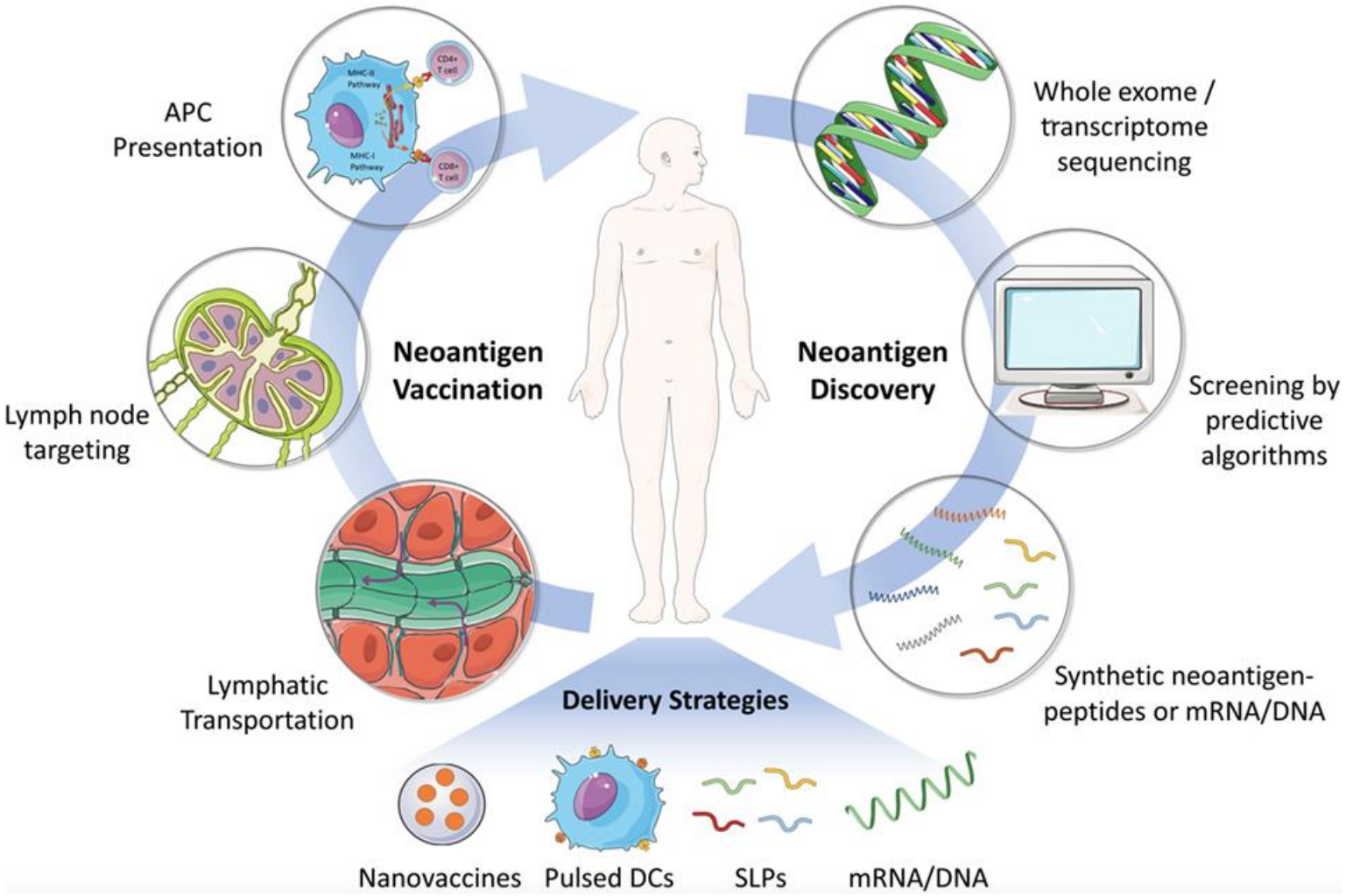
- QIAGEN CLC Genomics Workbench 20 Fusion Caller
  - Detects split-reads
  - Extra alignment step to artificially generated reference
    - Increase in the number of reads confidently mapped across fusion junctions

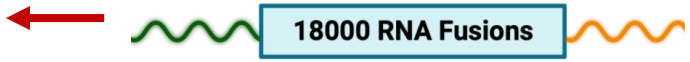
Overlap with TCGA Actionable Fusion Genes

CLC: 12 Actionable Fusions

Dragen/EasyFuse/Arriba: 1 Actionable Fusion

Peptide antigen and mRNA Vaccines for the Treatment & Prevention of Cancer

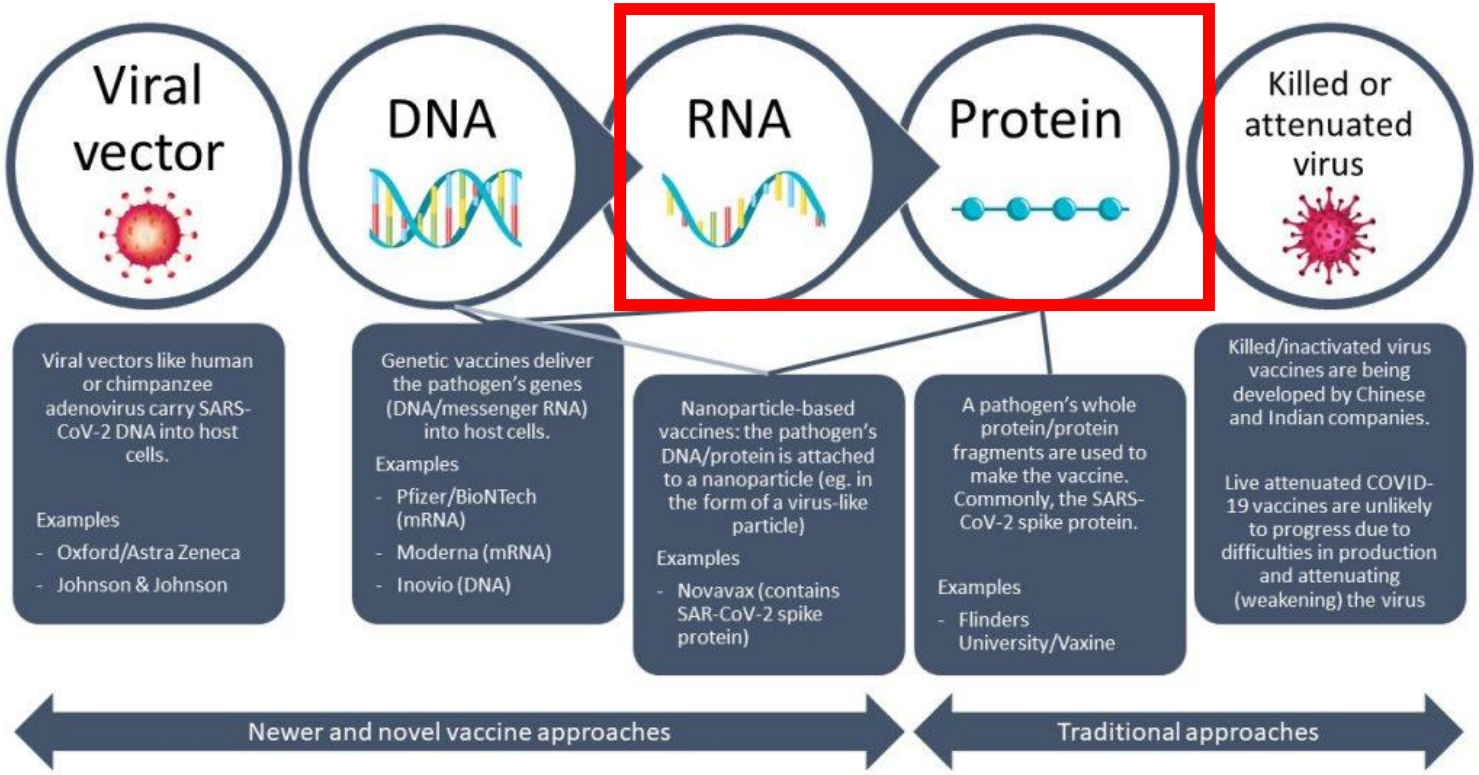




18,000 Chimeric RNAs identified from 1000 PDX models

Guo et al. 2018

Modern Approaches to Vaccination



Why Peptide Antigen Vaccines?

- Target very specific epitopes while minimizing the risks of allergic or autoimmune reactions

Our focus?

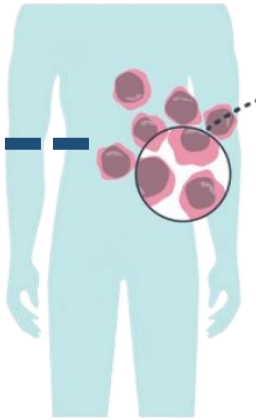
- Extracting neoantigens from fusion transcripts from two separate genes
- The unique sequences at the fusion junctions form new open reading frames (ORFs)

500 PDX Models

28 Cancer Types

18000 RNA Fusions

xenoSTART



Patient tissue blocks

1000 PDX
(Patient Derived Xenograft)
Tissue samples

UHSEQ
Sequencing Core

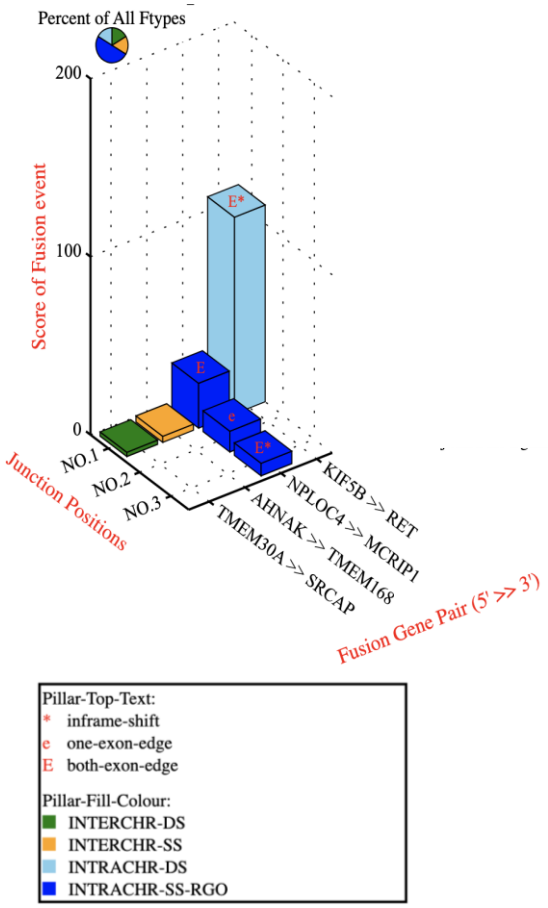
mRNA sequencing

Whole Exome Seq

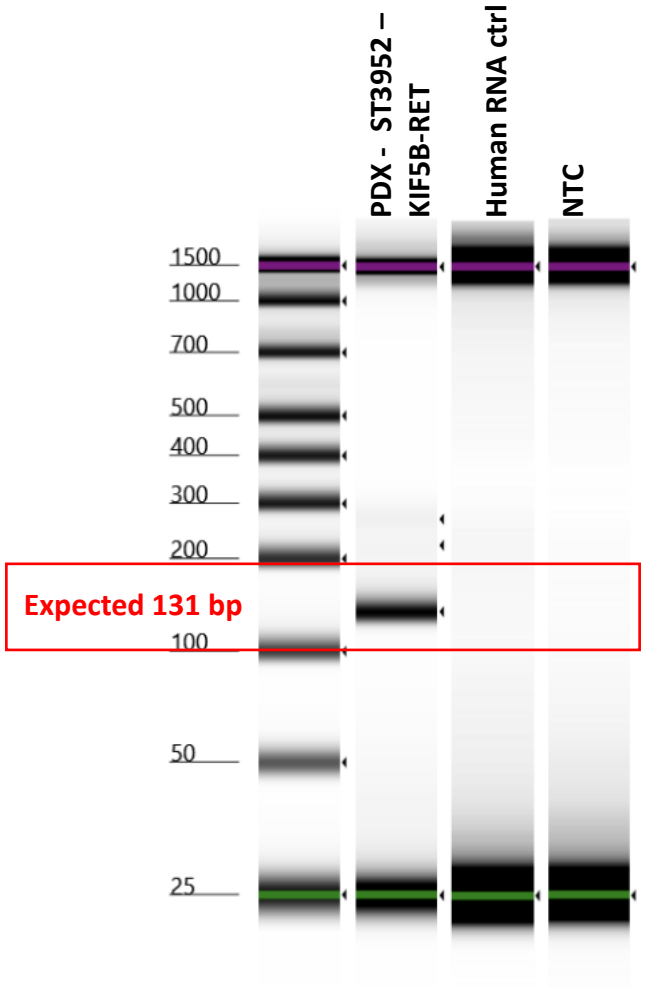
Neoantigen
Prediction
Fusion Analysis

RT-PCR and Sanger Validation of PDX Fusion KIF5B-RET

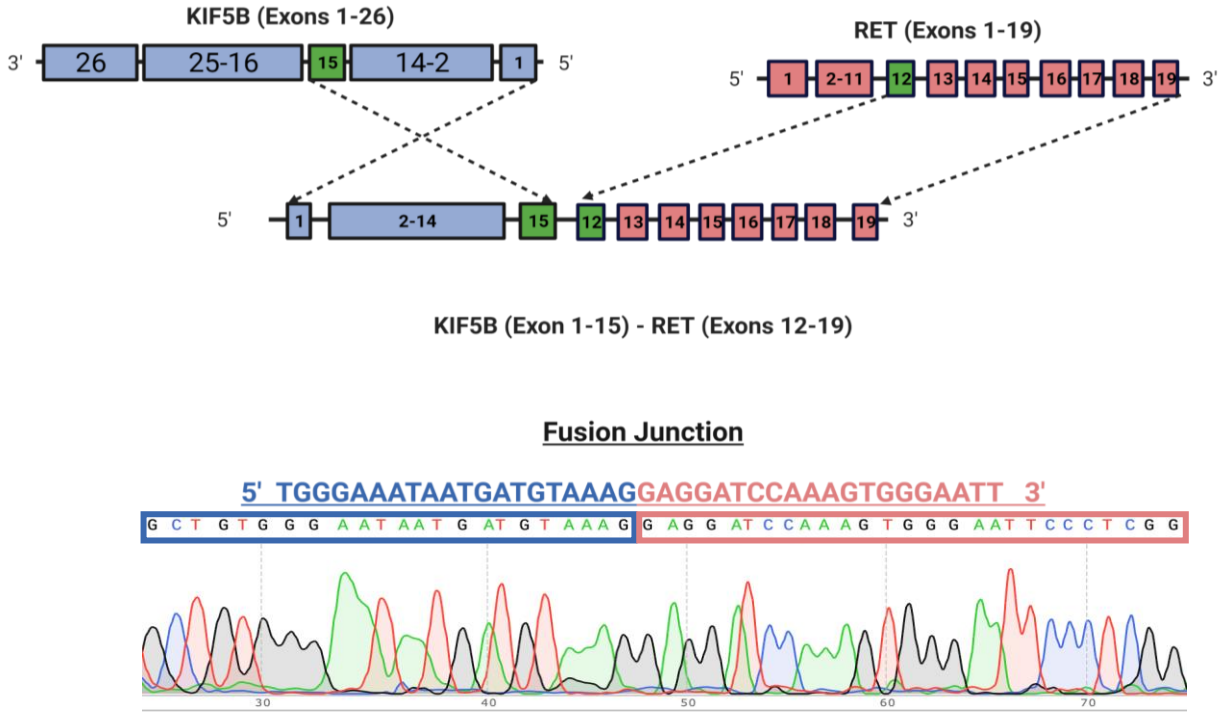
In-silico Identification



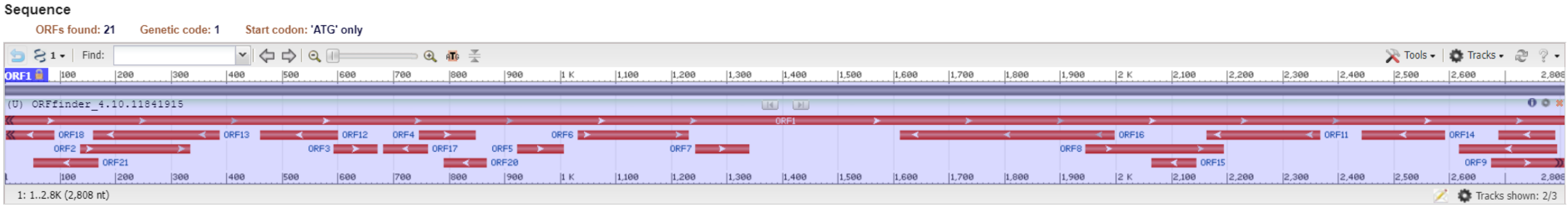
RT-PCR



Validation by Sanger Sequencing

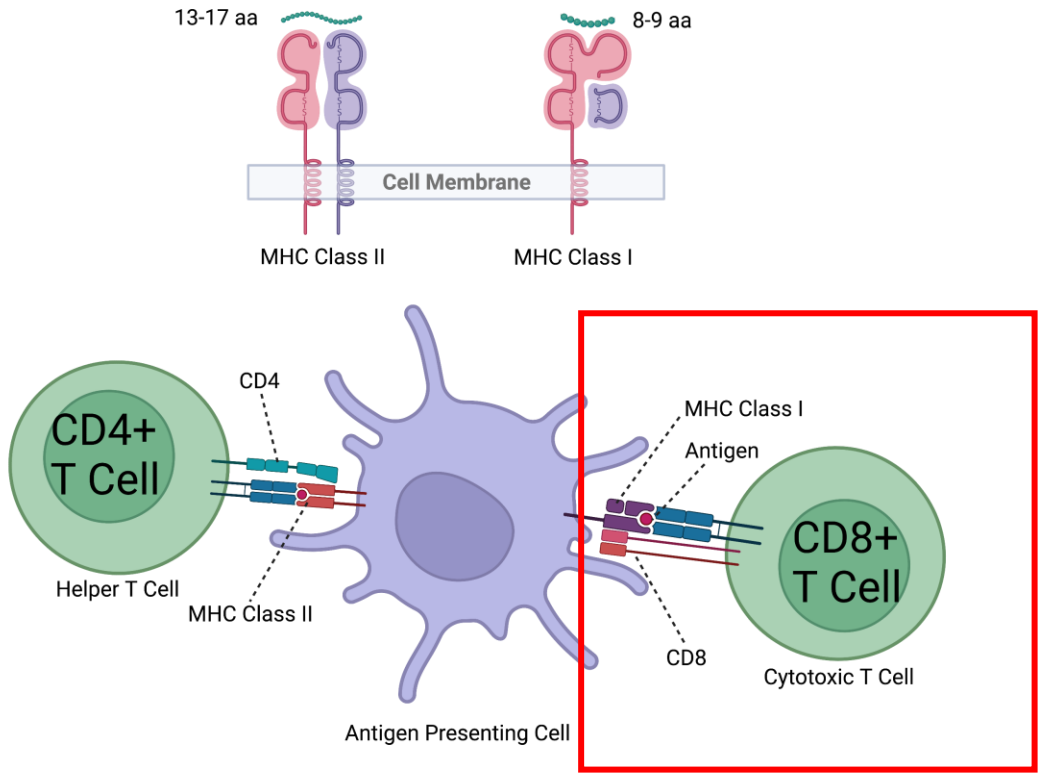


Neopeptide Region of KIF5B-RET Fusion

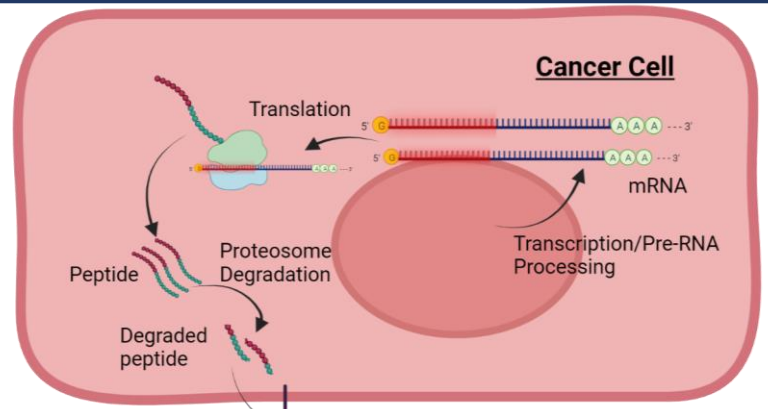


- 9-mer peptide sequences designed for peptide vaccines
- Yellow: Amino acid contribution from KIF5B
- Green: Amino acid contribution from RET

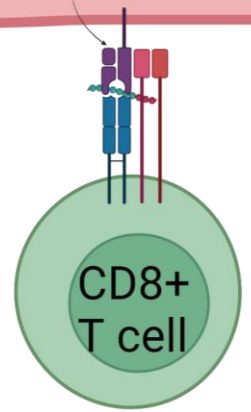
Fusion Presentation to CD8+ and CD4+ Requires MHC-Class I & II



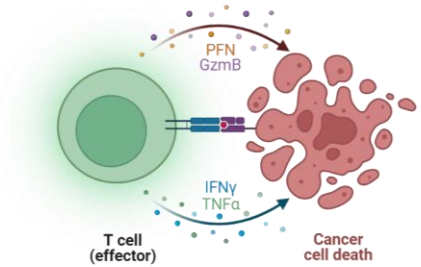
- CD8+ (cytotoxic) T cells directly kill cells presenting non-self epitopes.
- CD4+ (helper) T cells regulate the immune response to a particular antigen.



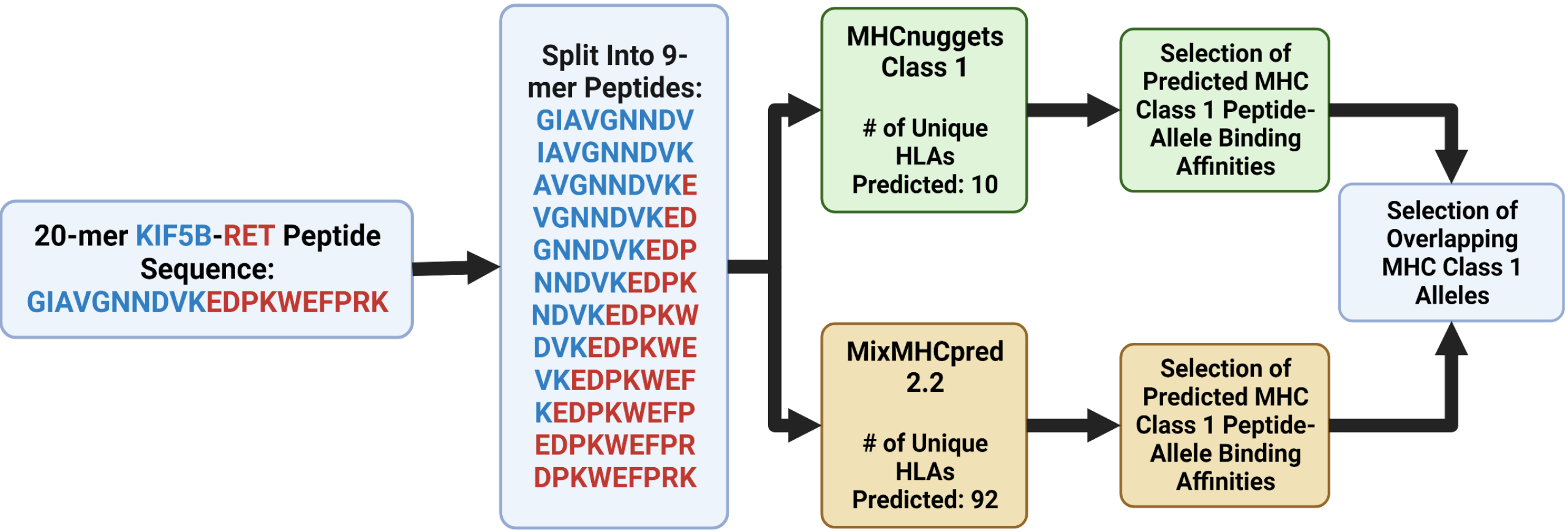
- Degraded proteins may be presented on MHC (Human Leukocyte Antigen - HLA) Class I molecules
  - Dependent on affinity
  - Non-self recognition by circulating CD8+ T cells



Non-Self Recognition - Cytotoxic Response



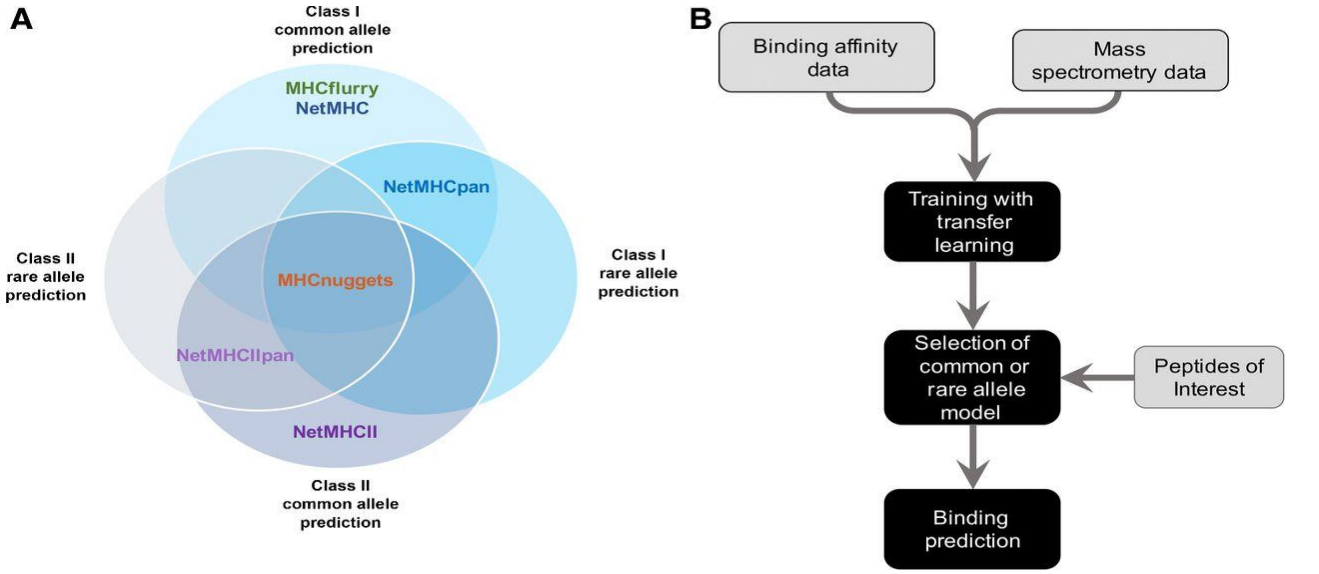
Neopeptide Affinity Prediction Pipeline



Neopeptide Affinity Prediction Pipeline

High-Throughput Prediction of MHC Class I and II Neoantigens with MHCnuggets

Xiaoshan M. Shao, Rohit Bhattacharya, Justin Huang, I.K. Ashok Sivakumar, Collin Tokheim, Lily Zheng, Dylan Hirsch, Benjamin Kaminow, Ashton Omdahl, Maria Bonsack, Angelika B. Riemer, Victor E. Velculescu, Valsamo Anagnostou, Kymberleigh A. Pagel, and Rachel Karchin



Fragmented Peptides (9-mer) → **Binary Results table of binding affinity** → **Distribution of MHC alleles**

- **Deep-learning models**
- **Transfer-learning protocol**
  - Improves predictive power by circumventing issues due to variability of peptide length and the lack of data concerning rare HLA alleles.

Mei, *et al*., 2020 (Brief Bioinform)

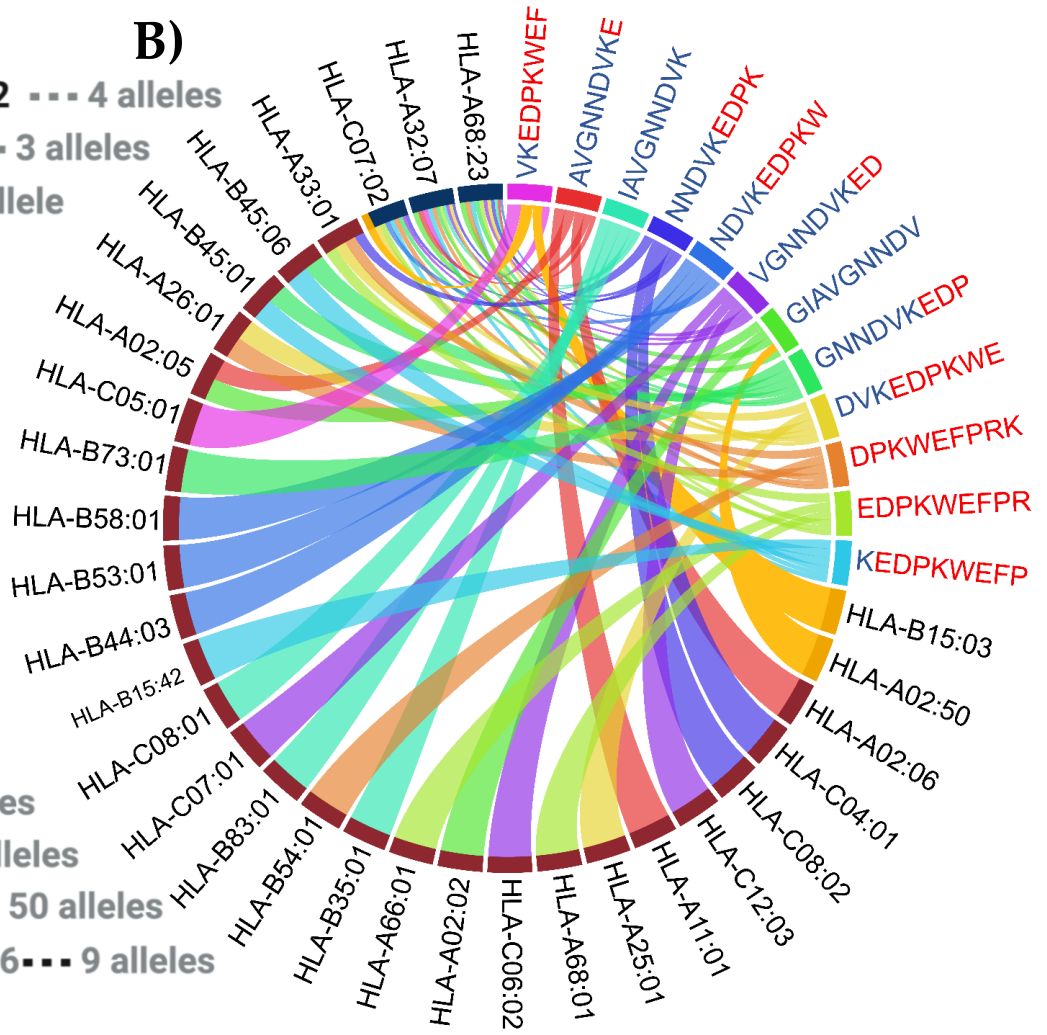
GfellerLab/ MixMHCpred

HLA-I ligand predictor

1 Contributor · 0 Issues · 20 Stars · 7 Forks

- **Uses sequence-scoring functions as basis of predictions.**
  - Calculates Position-Specific Scoring Matrix (PSSM) of all the inputted HLA alleles.
  - Outputs %Rank value indicative of HLA allotype's binding affinity to a specified peptide.
  - Sensitivity (Sn), Specificity (Sp), Accuracy (Acc)
- **Cutoff value of 10.00 for %Rank values was considered a strong binder**
  - HLA alleles that fell within this cutoff value often had IC50 values less than 500nM when compared to results from MHCnuggets.

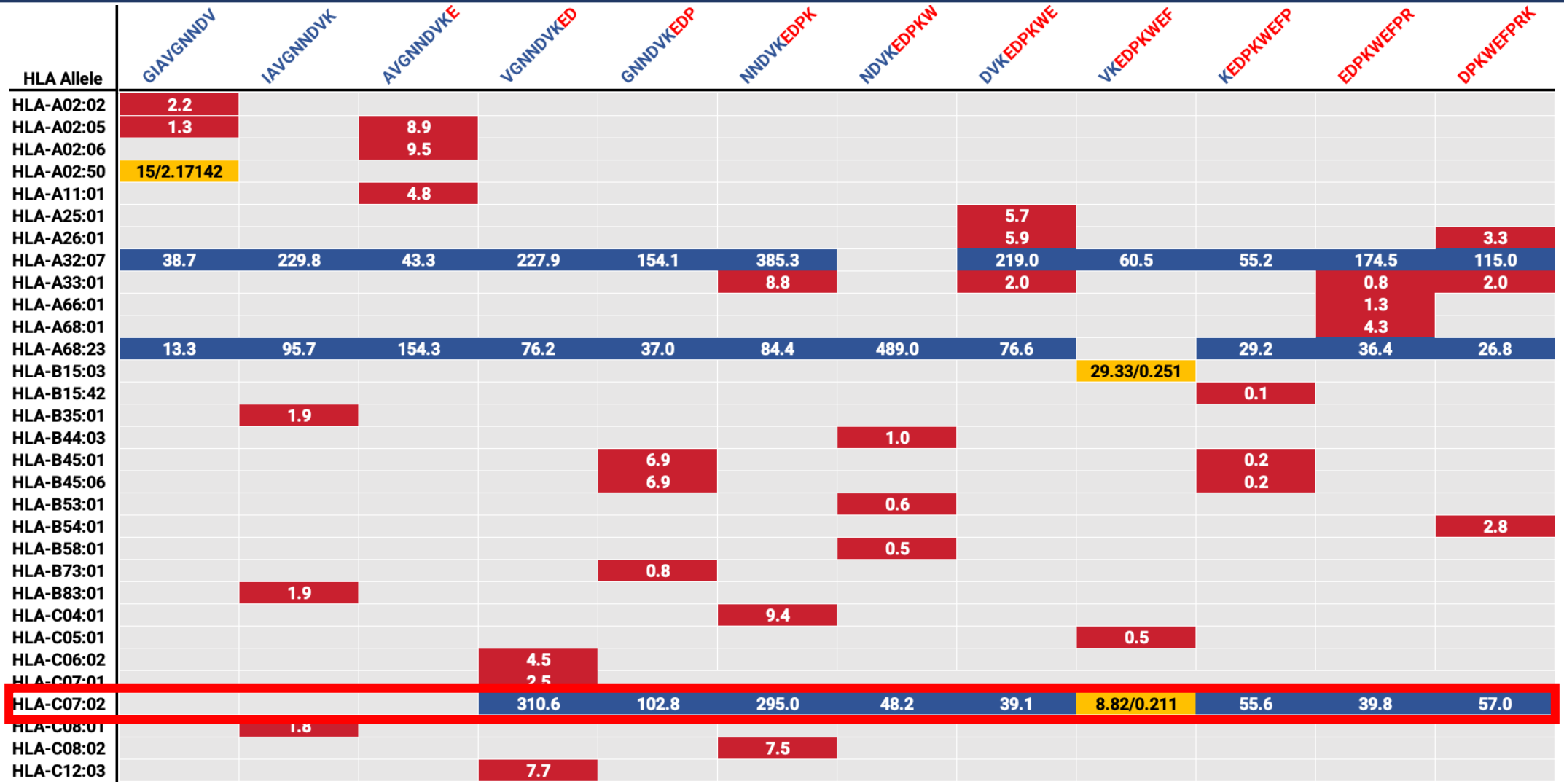
Neopeptide Affinity Predictions



A) Binding of junction peptides to MHC Class 1 alleles. Lines in grey represent calls made by MHCNuggets. Lines in **black** represent calls made by MixMHCPred.

B) Circos plot of neoantigen binding. **Red** and **blue** origins denote HLAs called by MixMHCPred and MHCNuggets, respectively. Gold chords denote overlapping calls.

HLA-C07:02 Shows Strong Binding to KIF5B-RET Neopeptides



HLA Allele	GIAVGNNDV	IAVGNNDVK	AVGNNDVKE	VGNNDVKED	GNNDVKEDP	NNDVKEDPK	NDVKEDPKW	DVKEDPKWE	VKEDPKWEF	KEDPKWEFP	EDPKWEFPR	DPKWEFPRK
HLA-A02:02	2.2											
HLA-A02:05	1.3		8.9									
HLA-A02:06			9.5									
HLA-A02:50	15/2.17142											
HLA-A11:01			4.8									
HLA-A25:01								5.7				
HLA-A26:01								5.9				3.3
HLA-A32:07	38.7	229.8	43.3	227.9	154.1	385.3		219.0	60.5	55.2	174.5	115.0
HLA-A33:01						8.8		2.0			0.8	2.0
HLA-A66:01											1.3	
HLA-A68:01											4.3	
HLA-A68:23	13.3	95.7	154.3	76.2	37.0	84.4	489.0	76.6		29.2	36.4	26.8
HLA-B15:03									29.33/0.251			
HLA-B15:42										0.1		
HLA-B35:01		1.9										
HLA-B44:03							1.0					
HLA-B45:01					6.9					0.2		
HLA-B45:06					6.9					0.2		
HLA-B53:01							0.6					
HLA-B54:01												2.8
HLA-B58:01							0.5					
HLA-B73:01					0.8							
HLA-B83:01		1.9										
HLA-C04:01						9.4						
HLA-C05:01									0.5			
HLA-C06:02				4.5								
HLA-C07:01				2.5								
HLA-C07:02				310.6	102.8	295.0	48.2	39.1	8.82/0.211	55.6	39.8	57.0
HLA-C08:01		1.8										
HLA-C08:02						7.5						
HLA-C12:03				7.7								

Table 1: Neoantigen affinity prediction table. Red and blue cells denote HLAs called by MixMHCPred and MHCNuggets, respectively. Gold cells denote overlapping calls.

Structure Based Binding Affinities and Peptide Cross-reactivity

- Geometry prediction for peptide-HLAs
- Binding energy prediction for peptides
- Virtual screening of HLA binders
- Customizable workflows for analysis

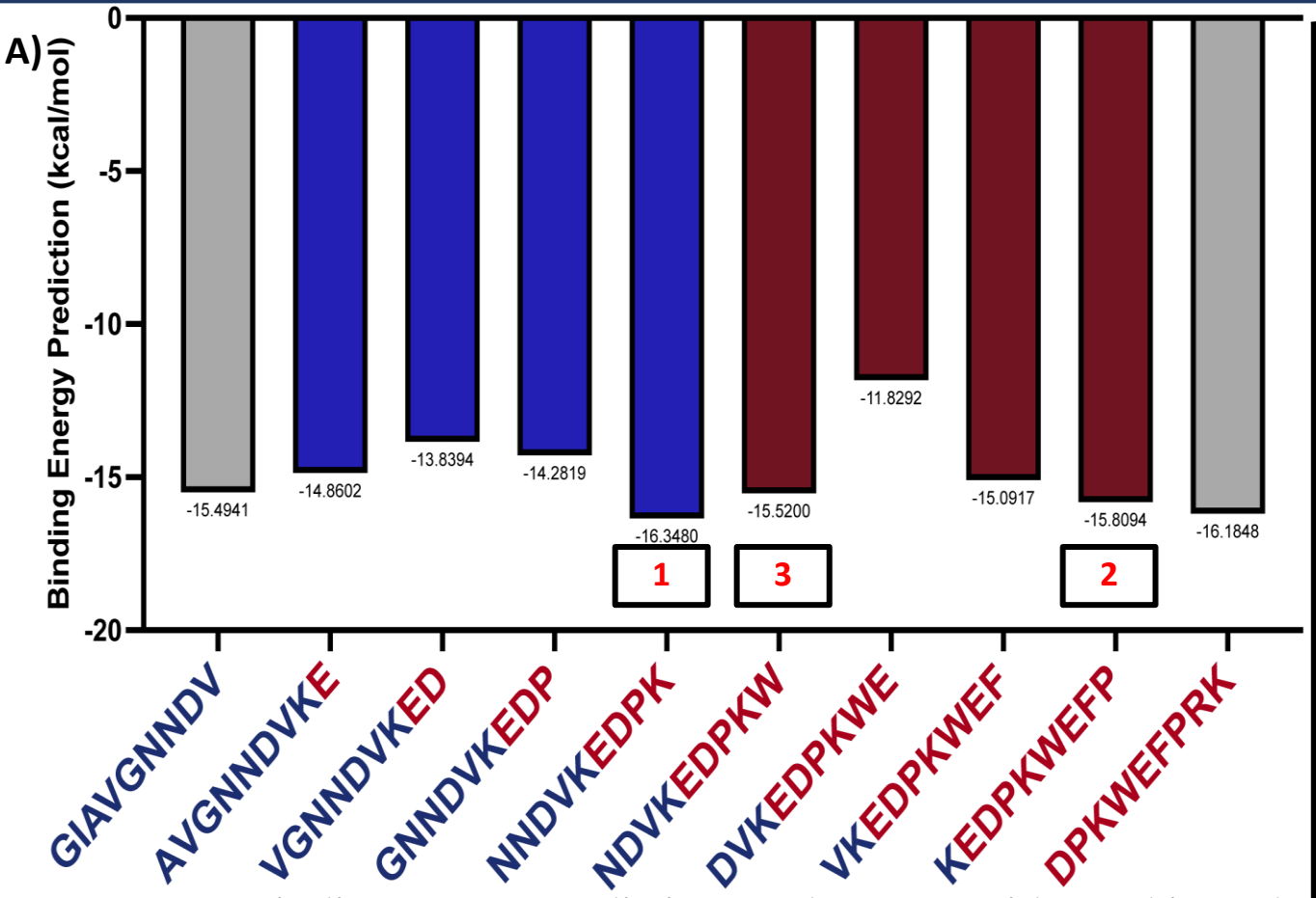
Considerations Taken

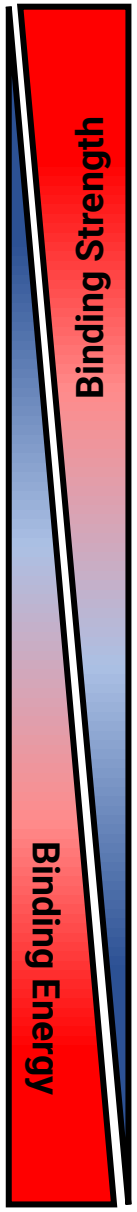
- **APE-GEN used to model 8 junction and 2 wild type peptides to the HLA-C*07:02 receptor.**
- **HLA-C*07:02 crystal structure 5VGE sourced from Protein Data Bank (PDB).**
- **R package Bio3d used to remove excess molecules and verify no chain breaks or gaps.**
- **Binding energies calculated with rigid structure.**

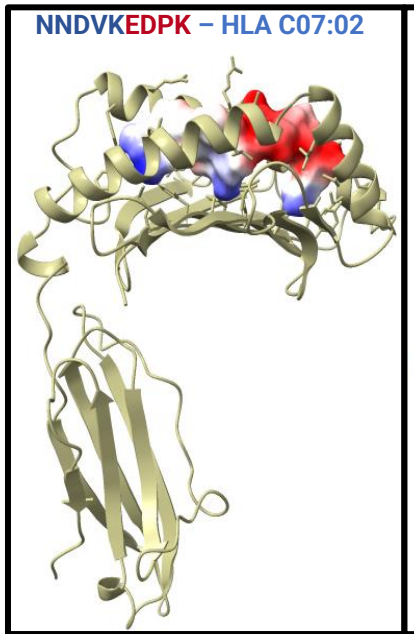
In Collaboration with: Jaila Lewis and Martiela Vaz de Freitas (Antunes Lab)

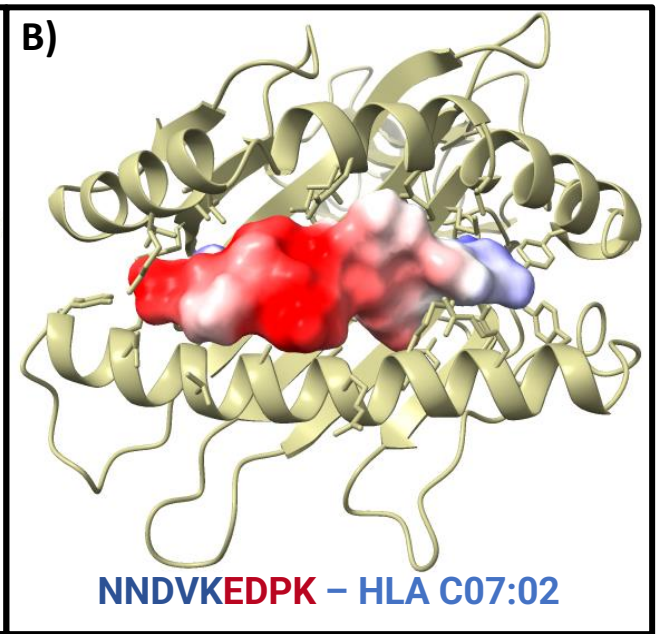
https://github.com/KavrakiLab/hla-arena
https://github.com/AntunesLab/crossdome

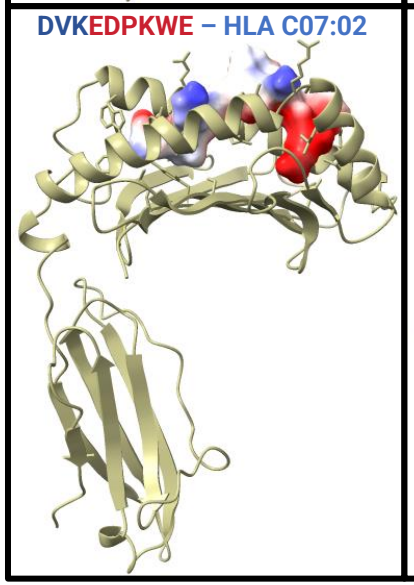
Structure Based Binding Affinity Predictions

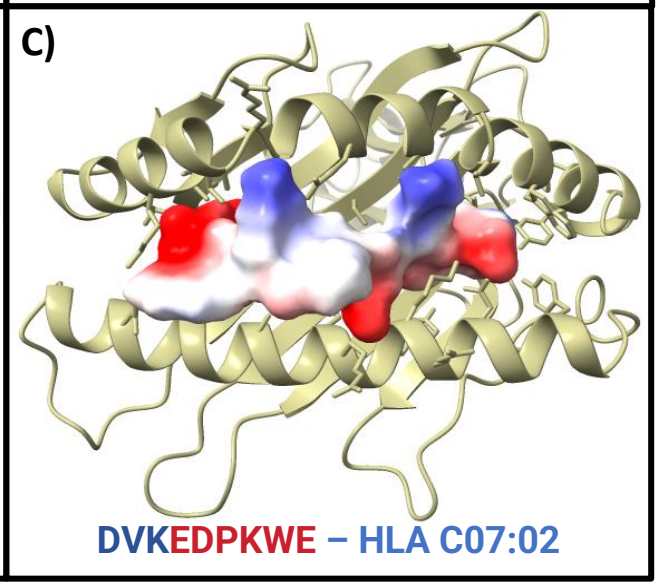












HLA-Arena Binding Energy Predictions and HLA-Peptide Docking. A) Binding energy predictions for junction peptides and two WT peptides (grey). Bars colored by majority contribution of amino acids in blue and red for KIF5B and RET respectively. **B)** and **C)** Peptides with highest and lowest average binding energy prediction values across three replicates. Peptides are docked in HLA-C07:02 and are colored by their electrostatic potential.

In Collaboration with: Jaila Lewis and Martiela Vaz de Freitas (Antunes Lab)

Structure Based Binding Affinities and Peptide Cross-reactivity

- Geometry prediction for peptide-HLAs
- Binding energy prediction for peptides
- Virtual screening of HLA binders
- Customizable workflows for analysis

Considerations Taken

- APE-GEN used to model 8 junction and 2 wild type peptides to the HLA-C*07:02 receptor.
- HLA-C*07:02 crystal structure 5VGE sourced from Protein Data Bank (PDB).
- R package Bio3d used to remove excess molecules and verify no chain breaks or gaps.
- Binding energies calculated with rigid

- Maps to Immunopeptidomic Data
- Identifies Self Derived Peptides
- Tissue Specific mRNA Expression

Considerations Taken

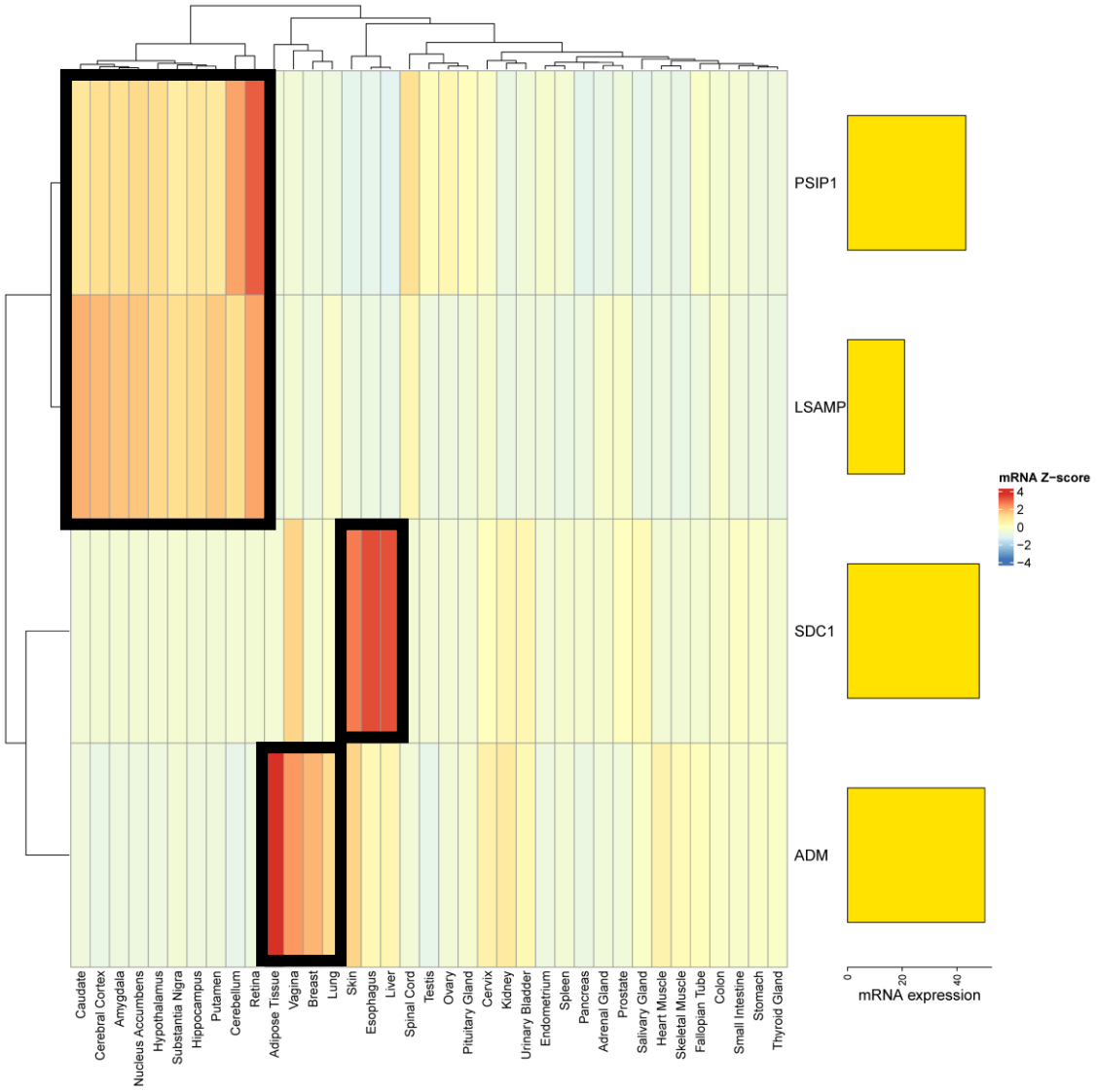
- 9-mers were limited to the HLA-C*07:02 allele to identify all potential self-derived peptides that may yield T-cell cross-reactivity.
- Default p-value cutoff for cross reactivity $p < 0.005$
  - Increased to <0.01 and <0.05 for this dataset

In Collaboration with: Jaila Lewis and Martiela Vaz de Freitas (Antunes Lab)

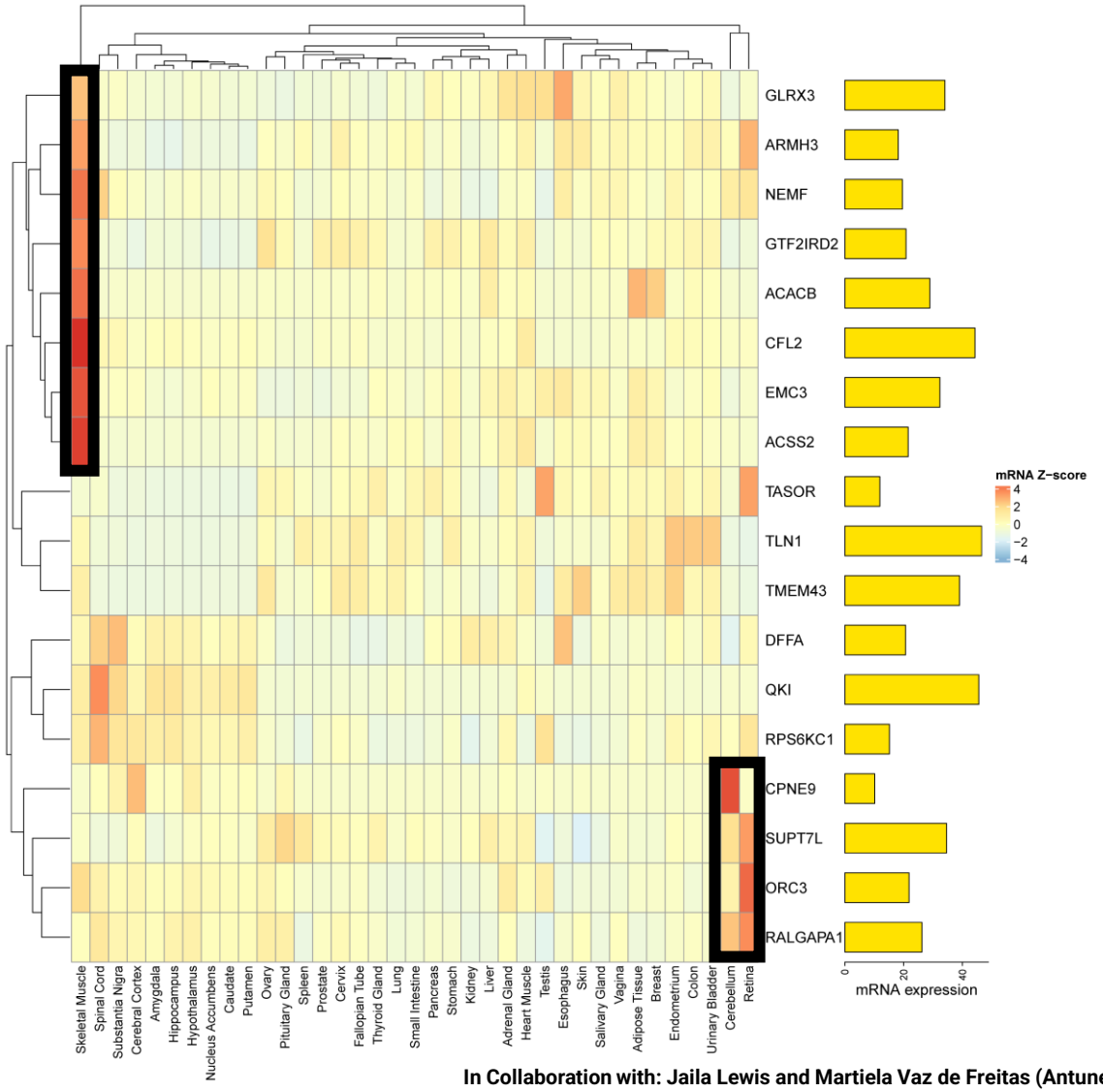
https://github.com/KavrakiLab/hla-arena
https://github.com/AntunesLab/crossdome

Assessment of T-Cell Cross Reactivity

NNDVKEDPK

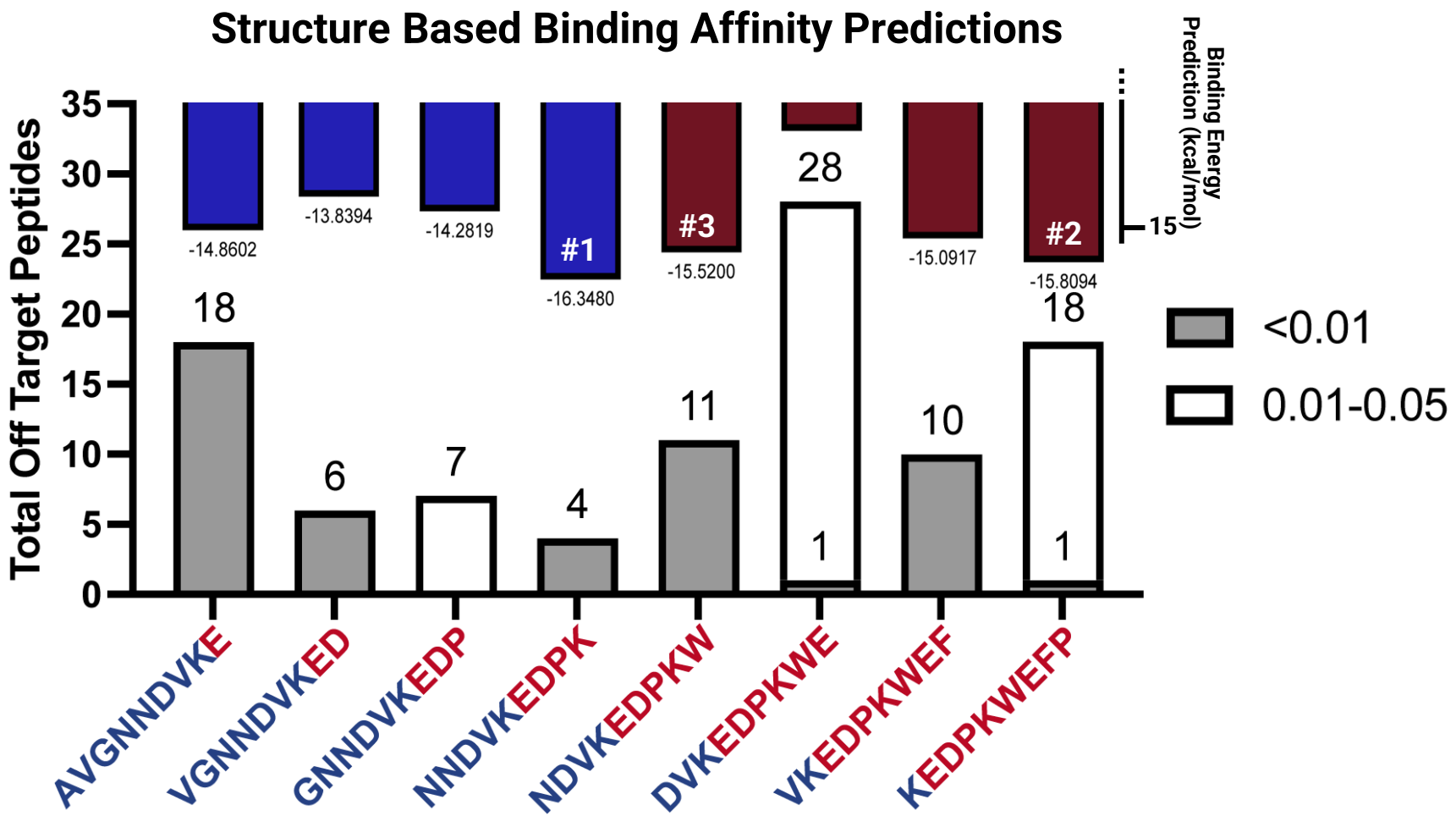


KEDPKWEFP



In Collaboration with: Jaila Lewis and Martiela Vaz de Freitas (Antunes Lab)

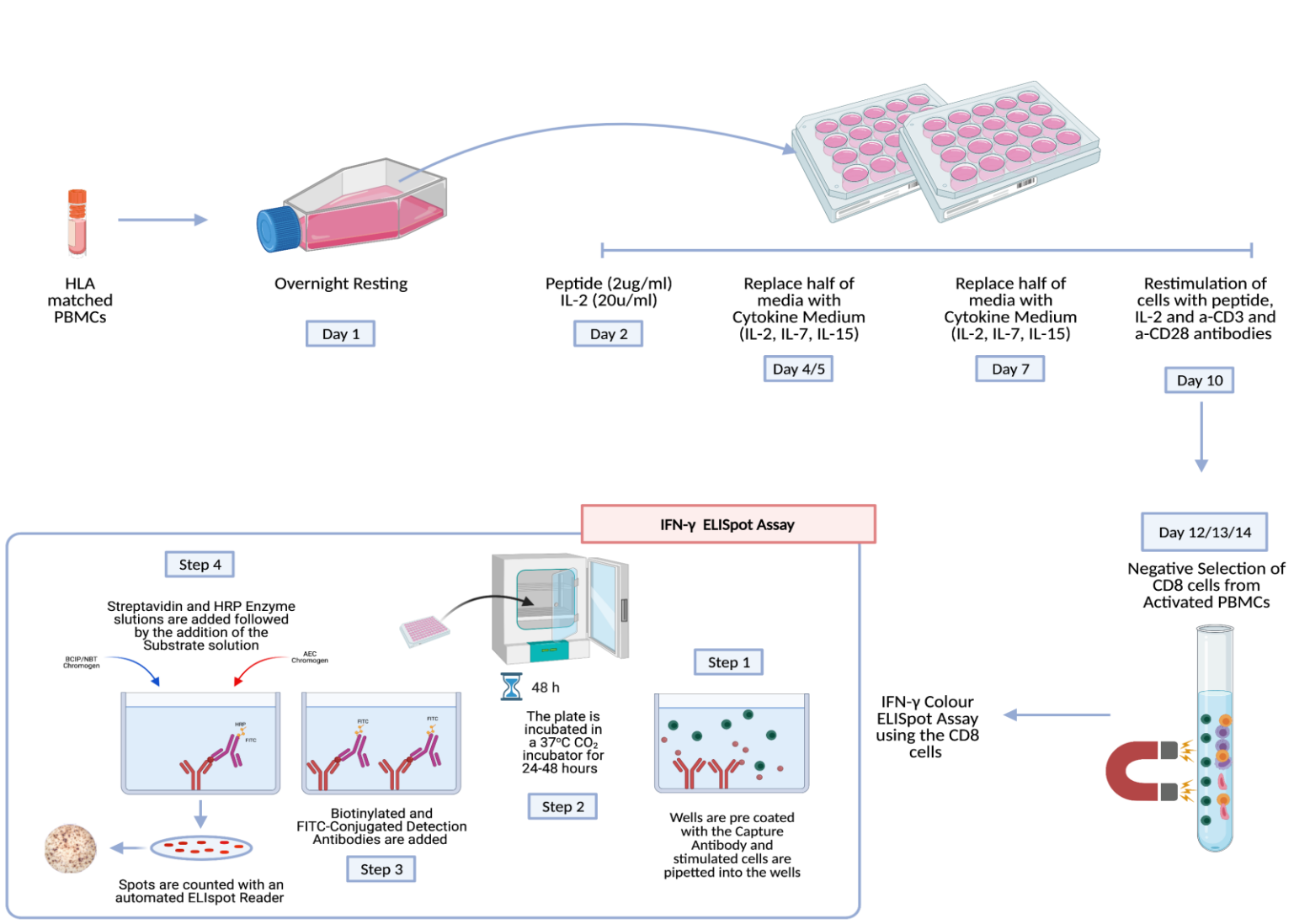
Crossdome Assessment of T-Cell Cross Reactivity



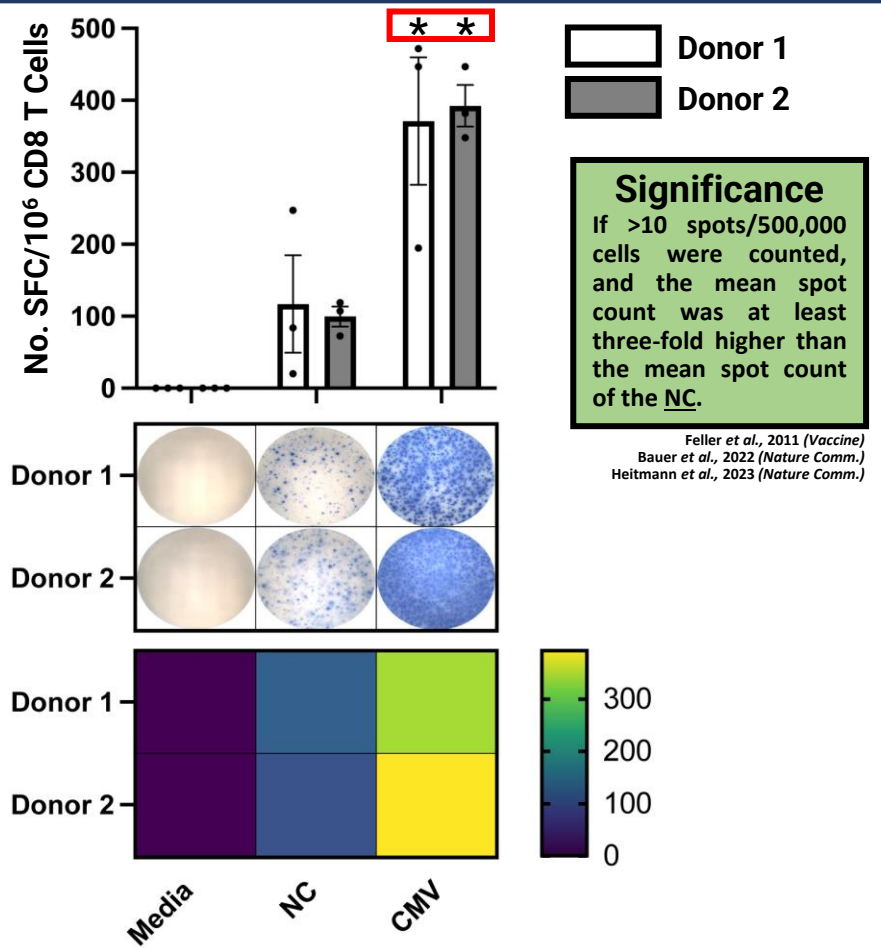
Top Ranked Off Target Neopeptides Per Neoantigen

In Collaboration with: Jaila Lewis and Martiela Vaz de Freitas (Antunes Lab)

Enzyme Linked Immunosorbent Spot Assay (ELISpot)



Created with BioRender



Significance

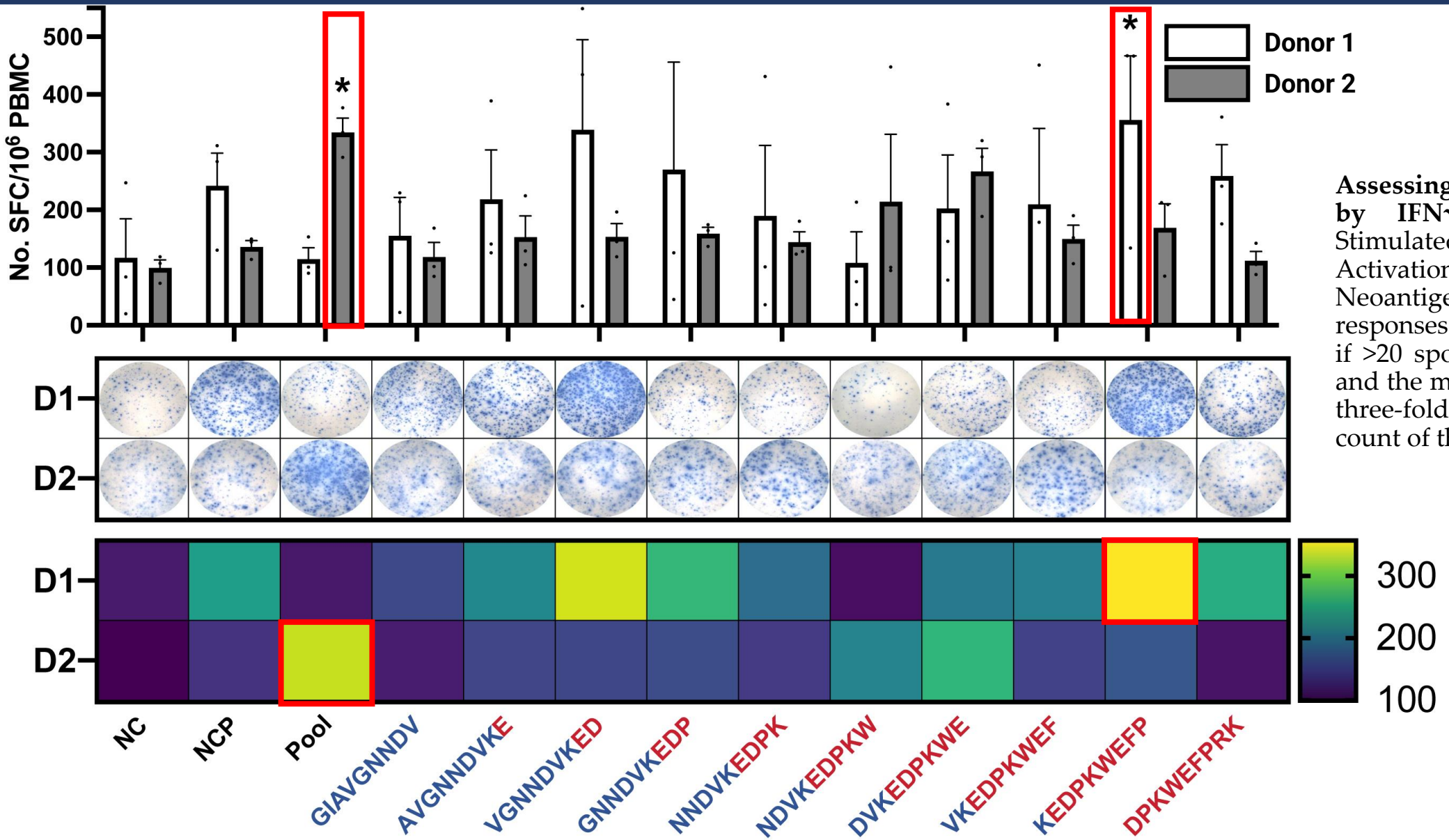
If >10 spots/500,000 cells were counted, and the mean spot count was at least three-fold higher than the mean spot count of the NC.

Feller *et al.*, 2011 *(Vaccine)*
Bauer *et al.*, 2022 *(Nature Comm.)*
Heitmann *et al.*, 2023 *(Nature Comm.)*

Assay Performance Controls

- **Negative Control**
  - Media Only
  - Non-Stimulated Control (**NC**)
- **Positive Control**
  - Cytomegalovirus (**CMV**) Peptide Pool Stimulation

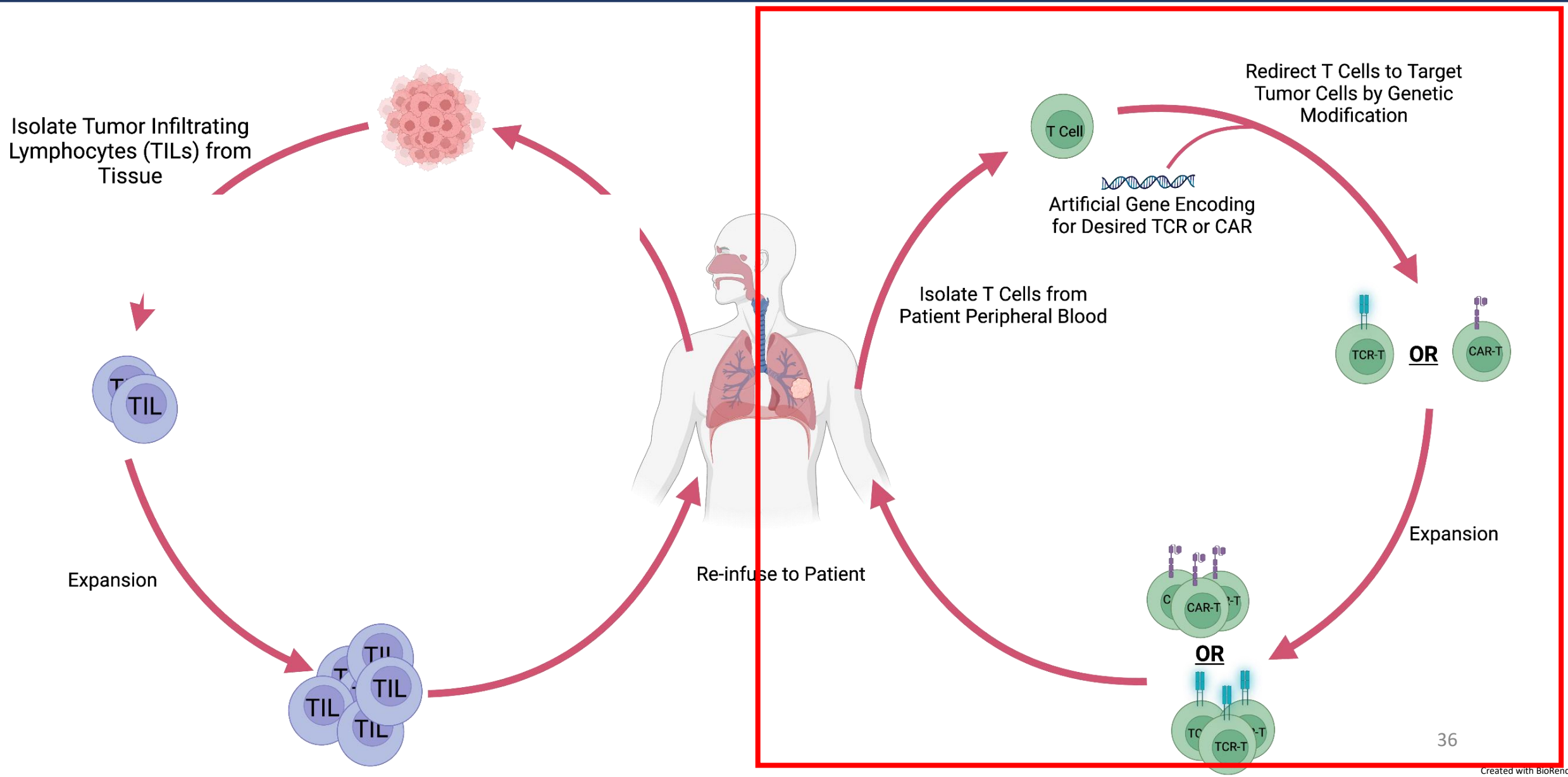
In-vitro Peptide Stimulation Reveals Responses Varying By Donor



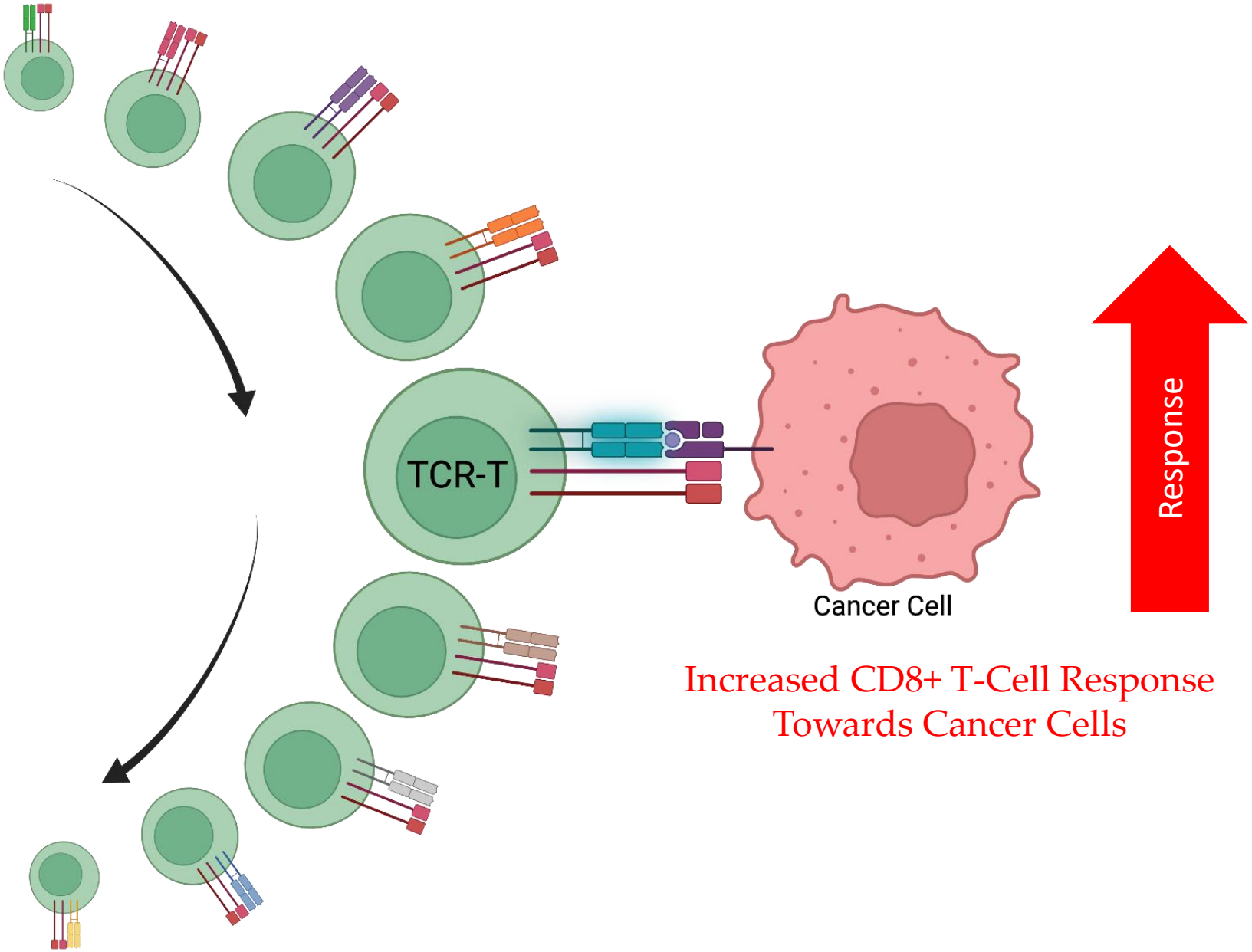
Assessing CD8+ T Cell Stimulation by IFNγ ELISpot. NC: Non-Stimulated Control, NCP: Non-Activation Control Peptide, NJP: Neoantigen Junction Pool. T cell responses were considered positive if >20 spots/1M cells were counted, and the mean spot count was at least three-fold higher than the mean spot count of the NC.

Castillo et al. (Unpublished)

Adoptive Cell Immune Therapies

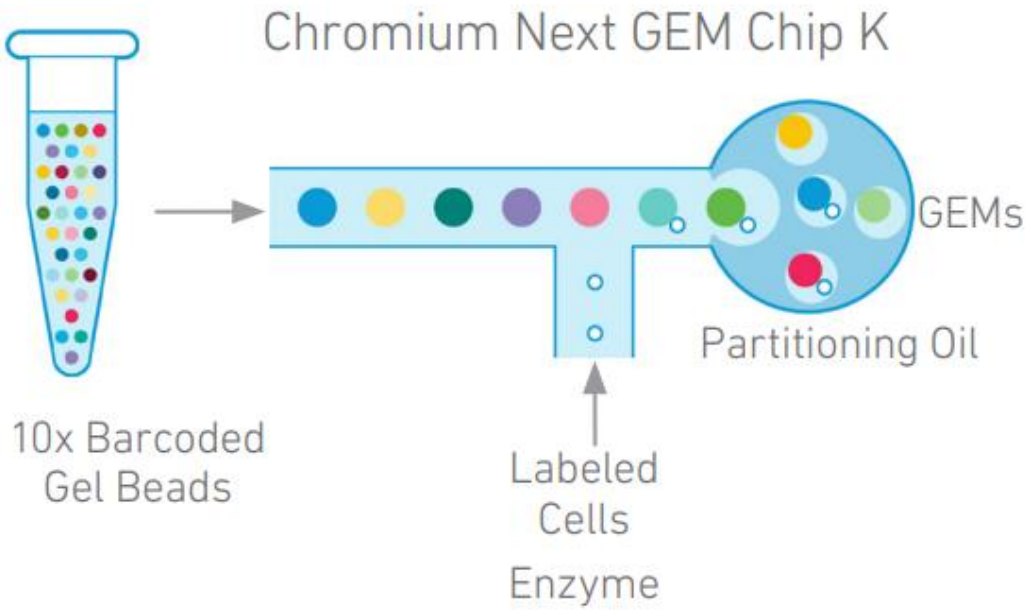


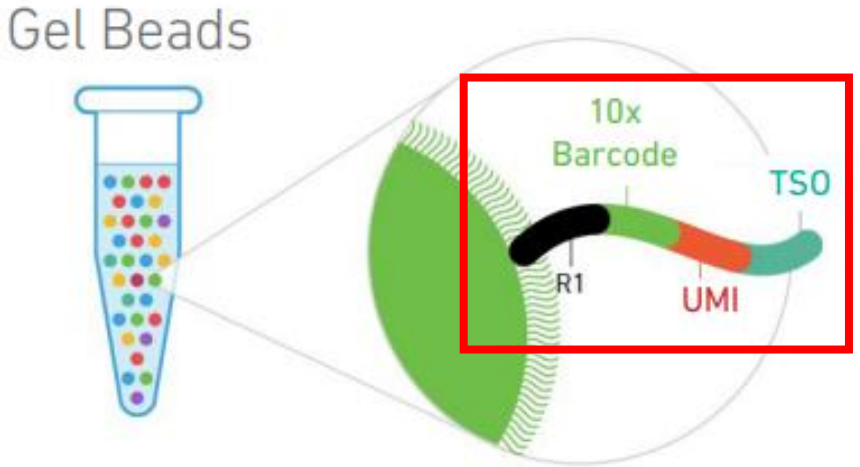
Identification of Expanded TCR Clonotypes in Peptide Stimulated T-Cells

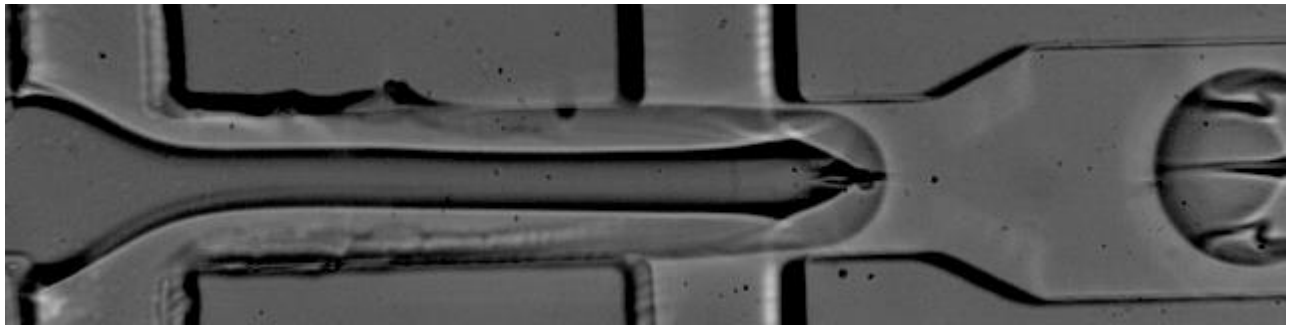


Increased CD8+ T-Cell Response
Towards Cancer Cells

10X Genomics Single Cell 5′ Gene Expression and Immune Profiling

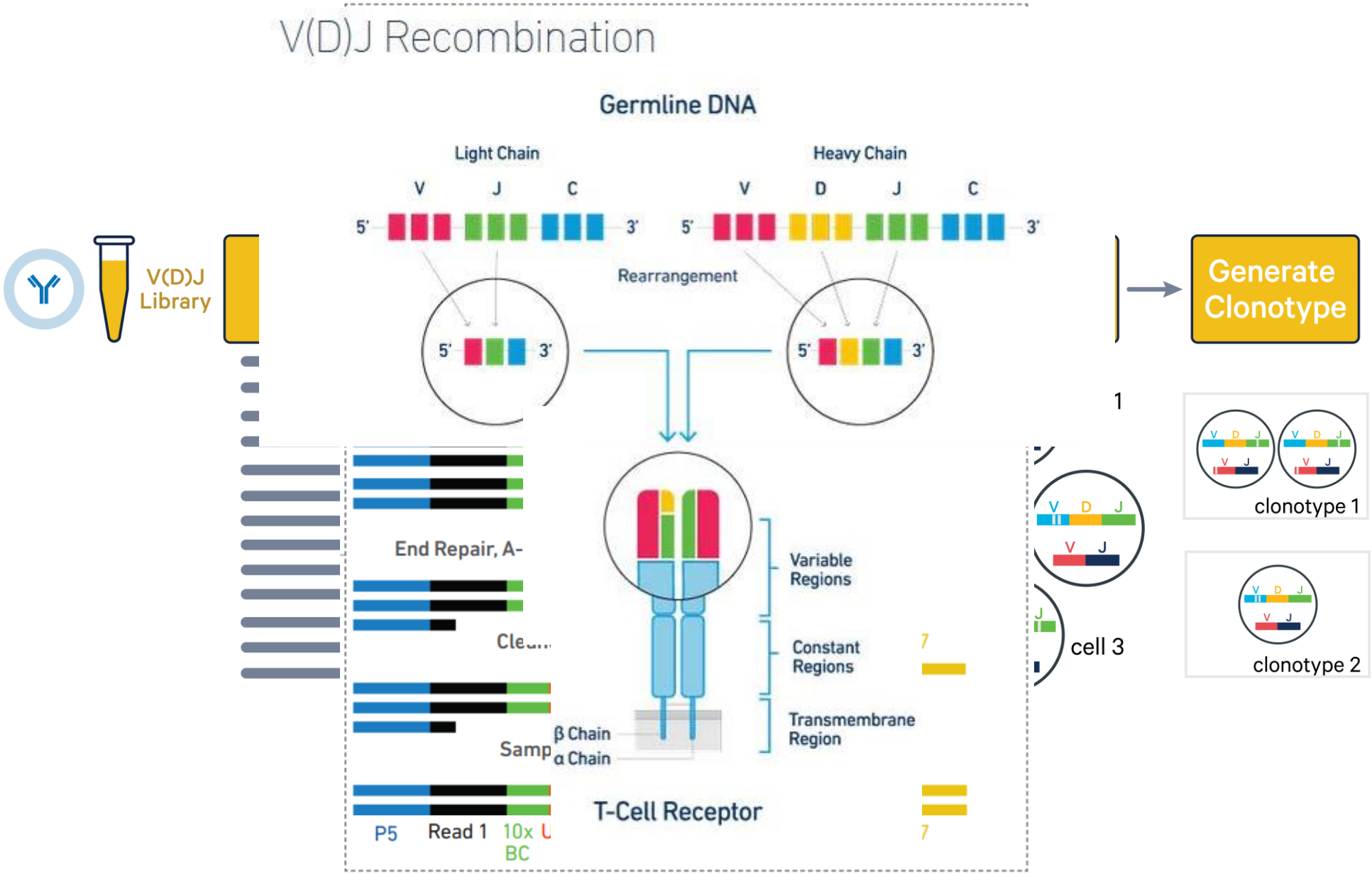






www.10xgenomics.com
https://mccarrolllab.org/dropseq/

VDJ Biology - Variability, (Diversity), Joining

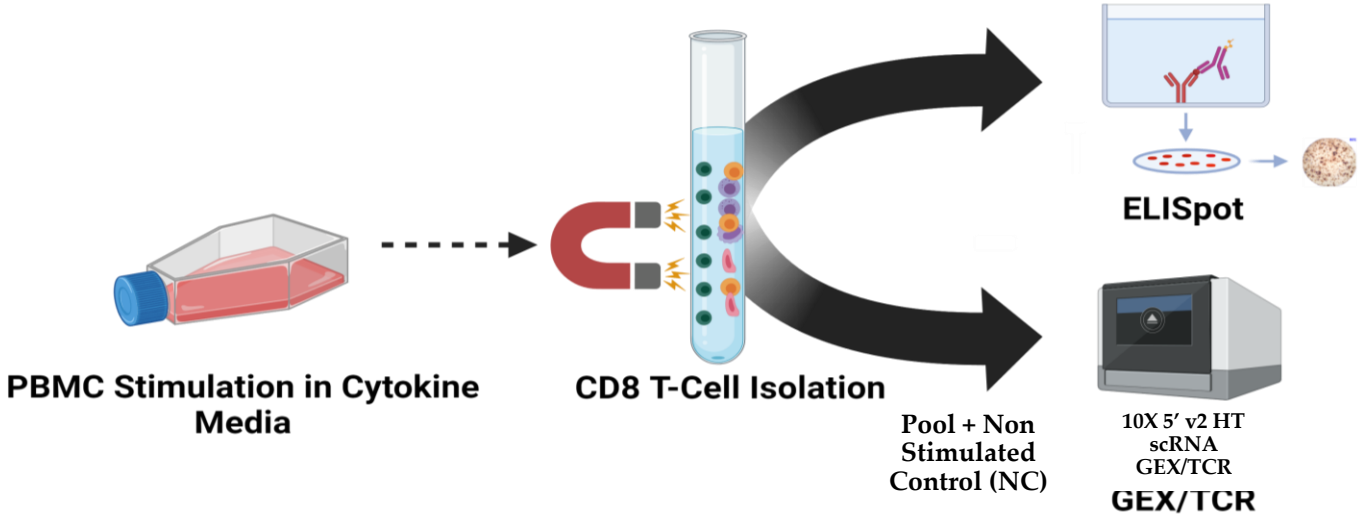


www.10xgenomics.com

Created with BioRender

Single Cell Immune Profiling: Sample Prep/Sequencing/Data QC

Sample Prep



Considerations Taken

- Cell Viability after CD8 T-Cell Isolation
  - 65-80% Live Cells
- Time between isolation and microfluidics chip loading
  - ~30 minutes
- Cells targeted
  - Donor 1: Lower cell counts elicited full sample loading
  - Donor 2: 12,000 cells

Sequencing + Data Analysis

NovaSeq X Plus
1 Lane per Donor

10X Genomics
Cloud Analysis

R Based:
Seurat
Data QC/Integration

Python Based:
scanpy:
Final Visualization
scirpy:
VDJ Analysis

Considerations Taken

- **NGS libraries pooled 80/20:GEX/TCR**
- **Cell Ranger Count/V(D)J v7.1.0**
- Doublets were removed using DoubletFinder and doublet rates specified by 10X 5′ HT protocol (UG:CG000423.RevC).
- Cells with high (>5%) mitochondrial counts or low features counts (<500) were removed.
- Seurat objects were converted to scanpy .h5ad files by converting metadata components into python (scripy) inputs.

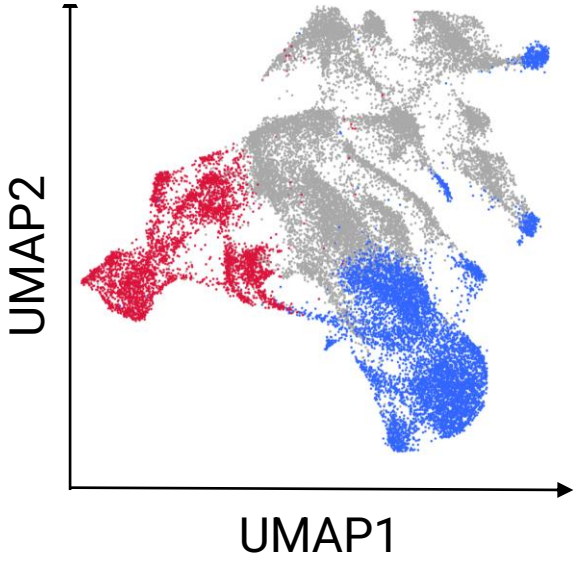
www.illumina.com, www.10xgenomics.com,https://satijalab.org/seurat/,https://scanpy.readthedocs.io/en/stable/,https://scirpy.scverse.org/en/latest/

Castillo et al. (Unpublished)

Activated, Exhausted, and Naïve T-Cell Populations Called Based on Canonical Markers

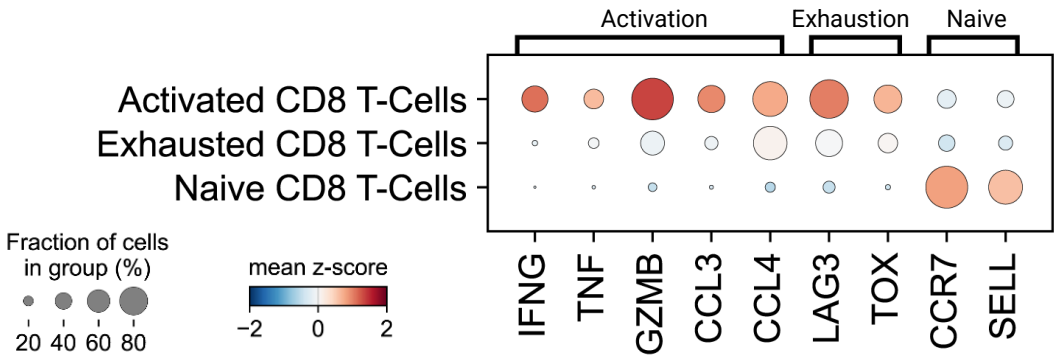
Defining CD8 T-Cell Populations

Isolated CD8 T-Cell Populations.



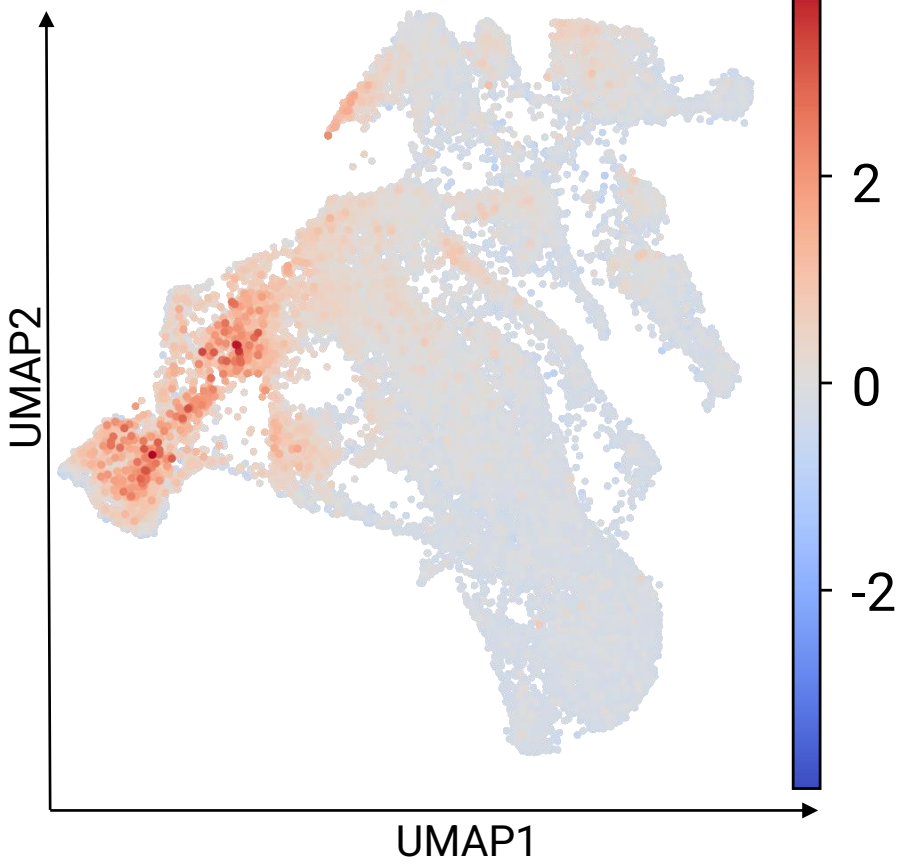
Marker Gene References
- Zhu *et al., 2020 (Cell)*
- Zheng *et al.*, 2021 (*Science*)
- Bauer *et al.*, 2022 *(Nature Comm.)*
- Chu *et al.*, 2022 *(Nature Med)*
- Tiethscher *et al., 2023 (Nature Comm.)*

- Activated CD8 T-Cells
- Exhausted CD8 T-Cells
- Naive CD8 T-Cells



Gene Scores Across Cell Populations

Activation Marker Score



Markers
IFNG
TNF
GZMB
CCL3
CCL4

Castillo et al. (Unpublished)

Characterization of Top Expanded TCR Clonotypes

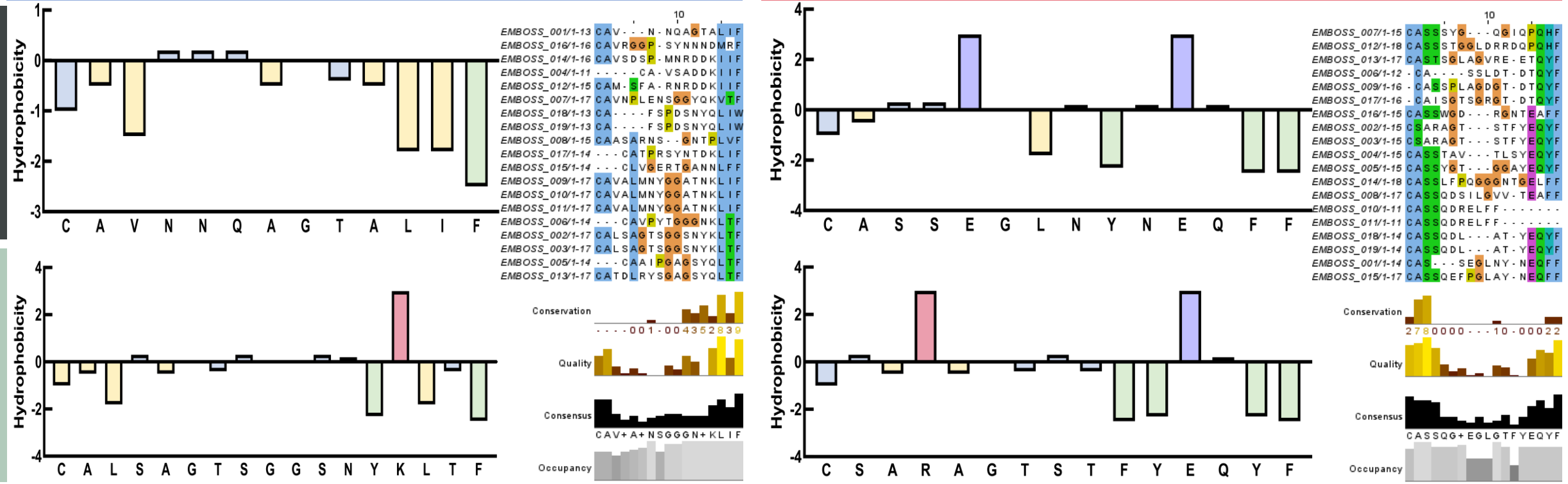
Amino Acid View of Top Expanded TCRs Per Donor

CDR3-α Chain

CDR3-β Chain

TCR 407

TCR 7939

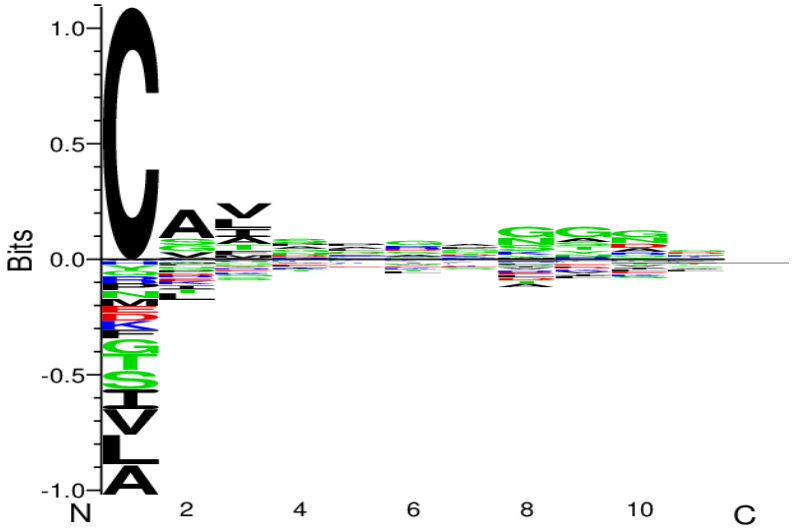
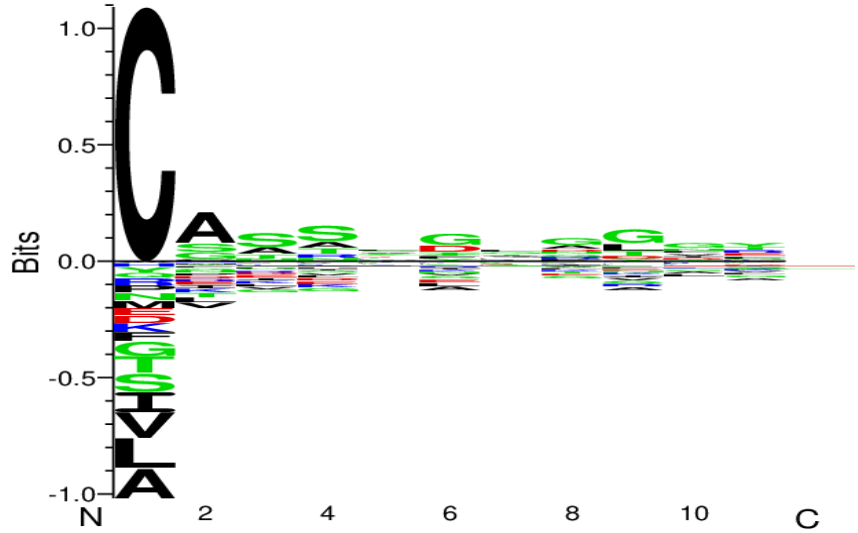
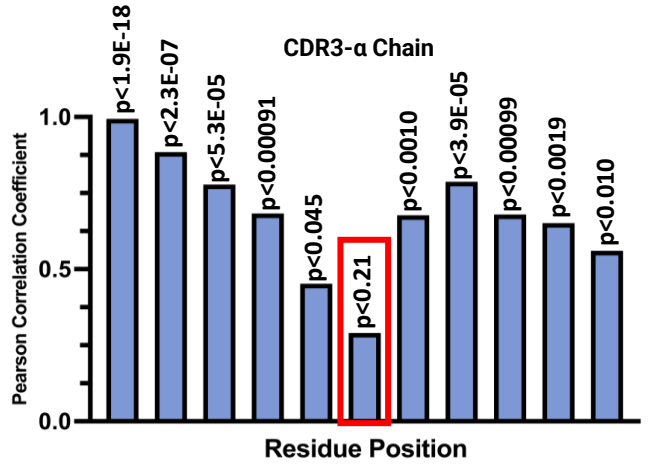
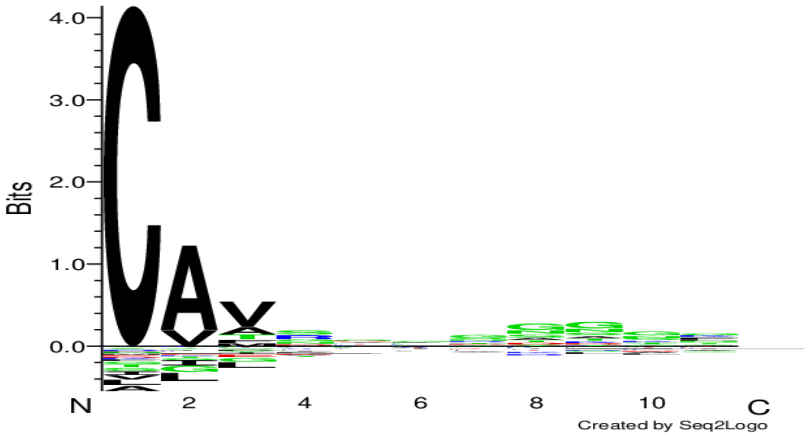
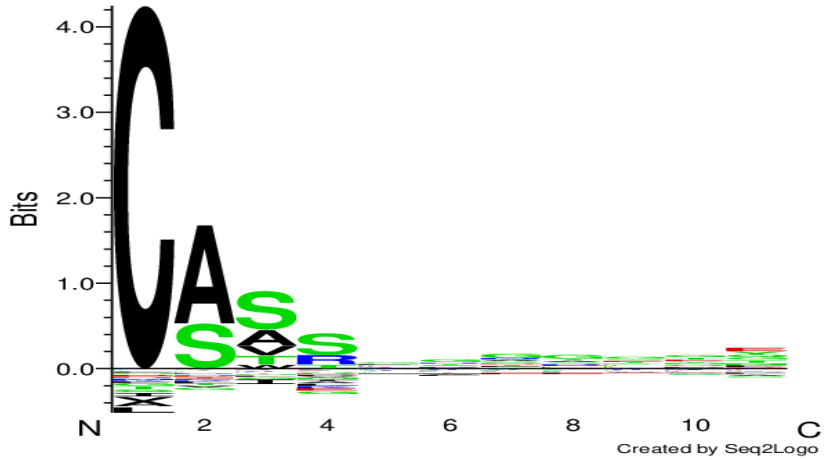
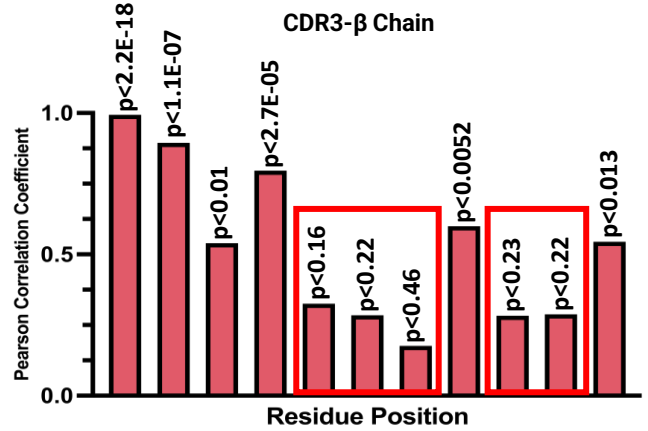


- AA Residues Colored by ClustalX Conservation

Clustal W: https://www.ebi.ac.uk/Tools/msa/clustalo/
Jalview v2.11.2.7

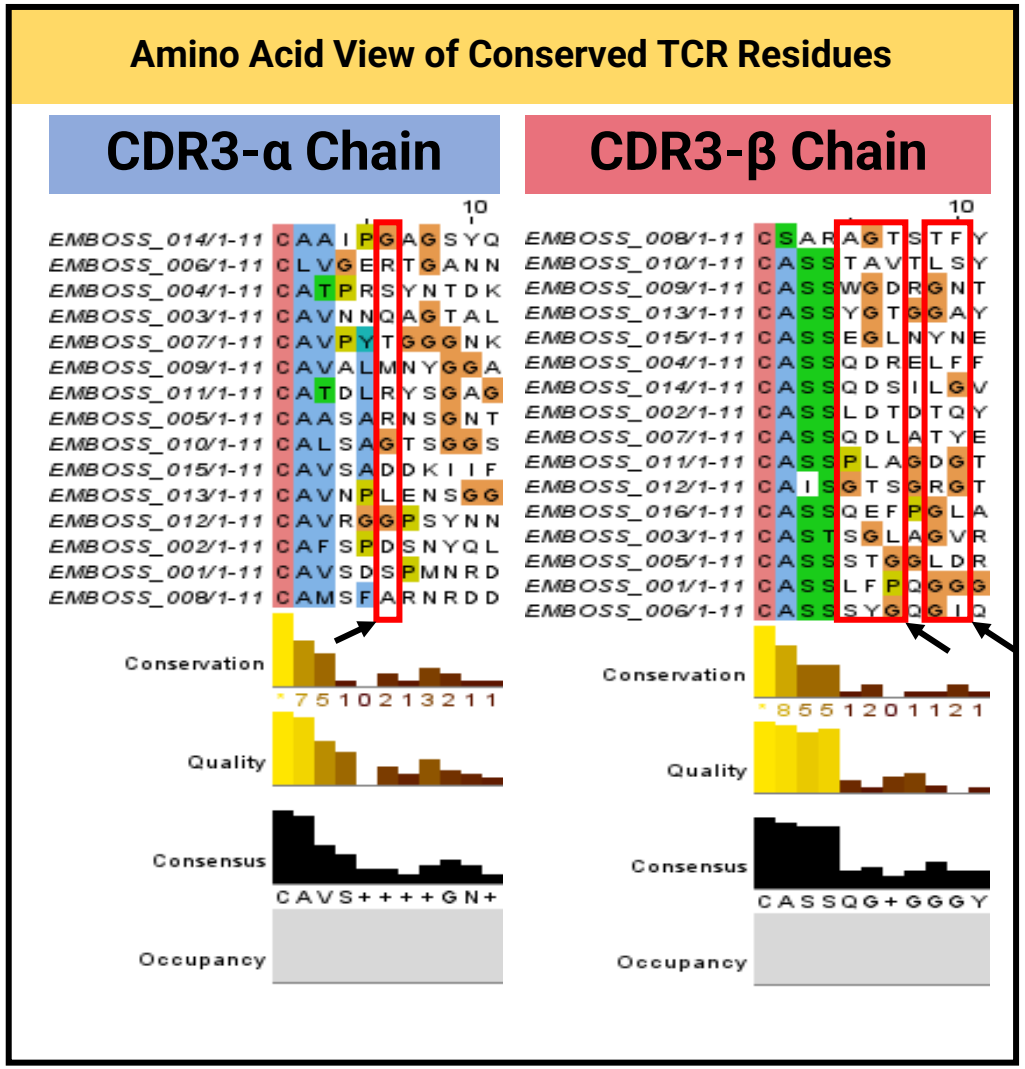
Identification of Expanded TCR Clonotypes in Peptide

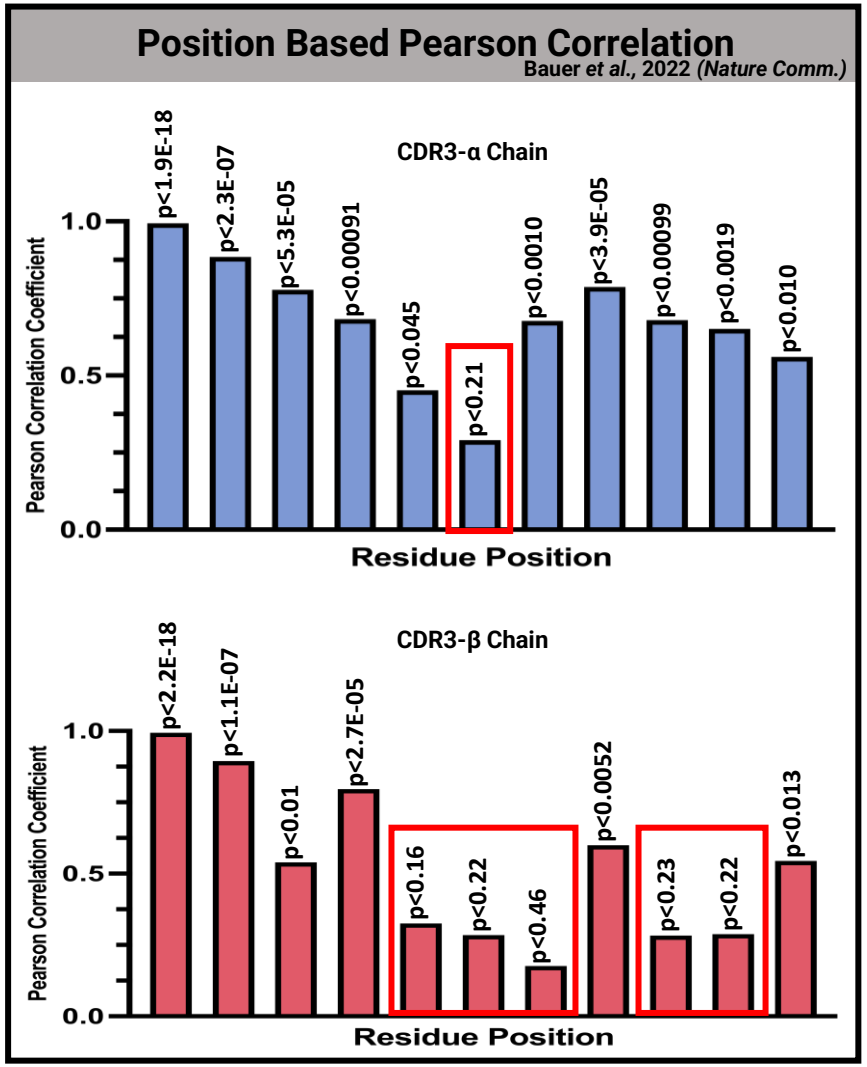
Variable Sequence Clustering, Motif Generation, and Residue Correlation

	CDR3-α Chain	CDR3-β Chain	Position Based Pearson Correlation (Bauer *et al.*, 2022 *(Nature Comm.)*)
Top Expanded TCRs			
Naïve TCRs			

https://services.healthtech.dtu.dk/services/GibbsCluster-2.0/ Castillo et al. (Unpublished)

Characterization of Top Expanded TCR Clonotypes

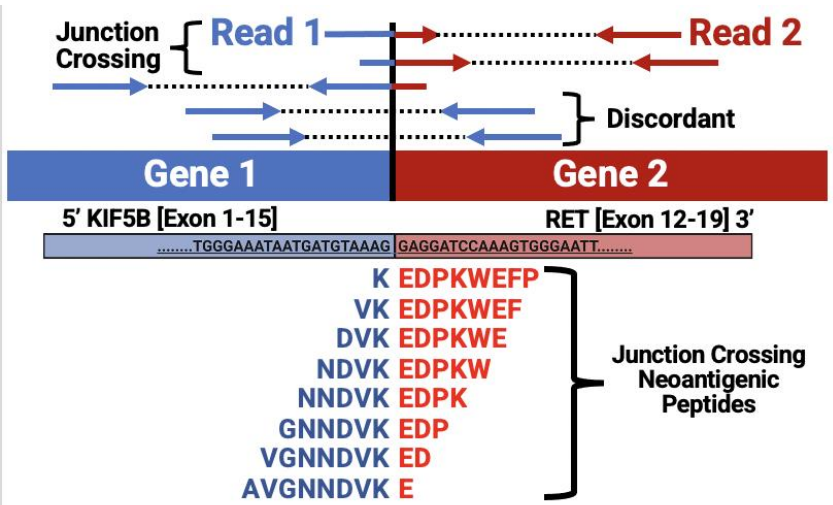


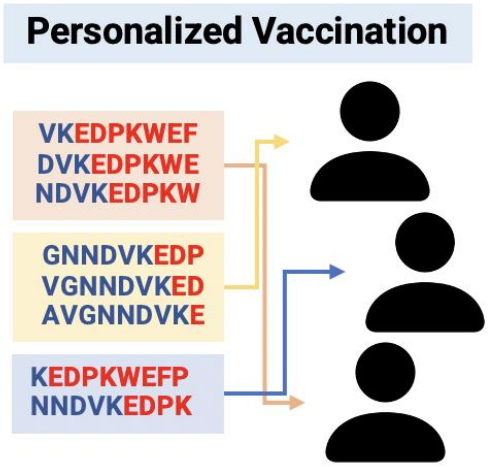


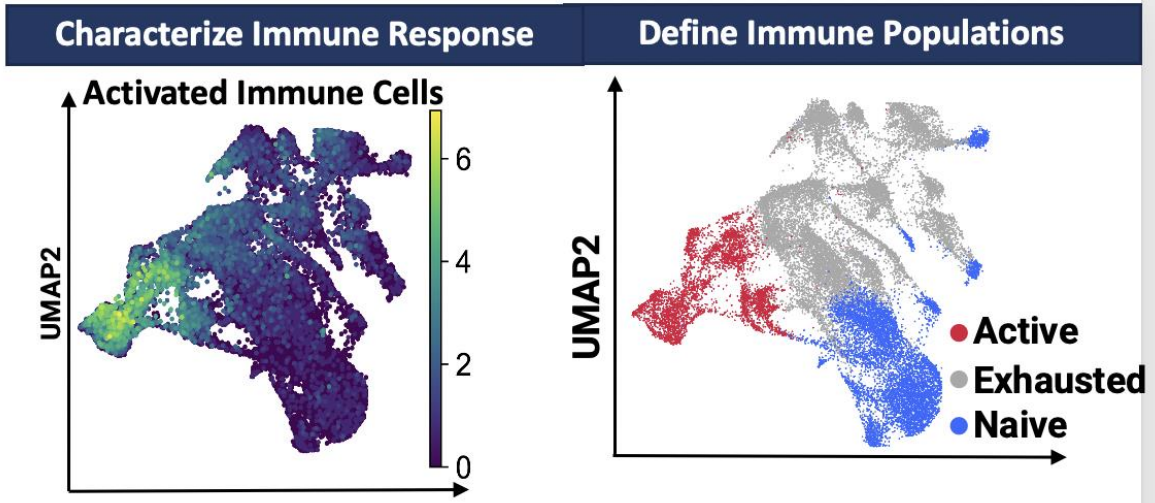
- AA Residues Colored by ClustalX Conservation

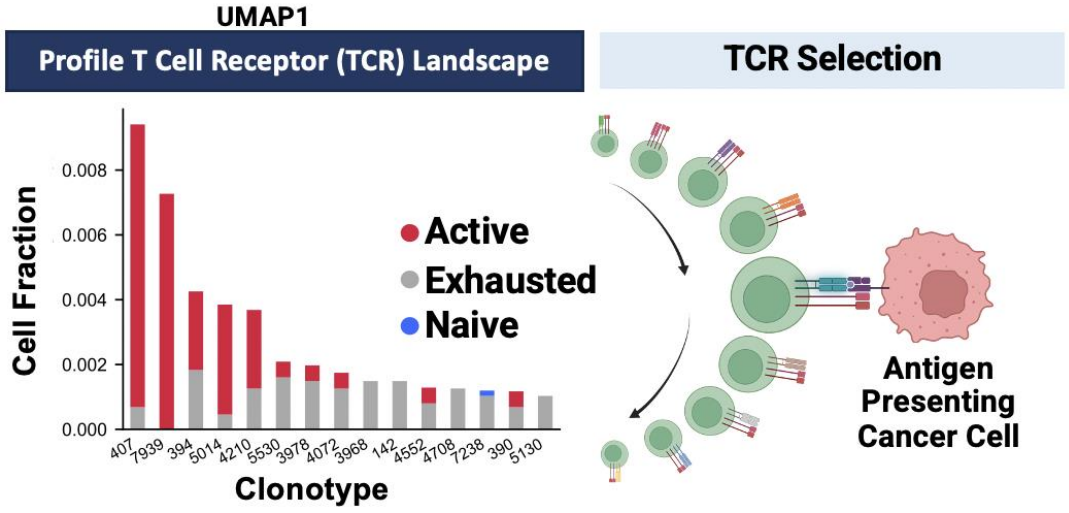
Variable Residue Glycines in Expanded T-Cells
- de Greef, *et al.*, 2021 (PNAS)
- Yassai, *et al.*, 2017 (*J. Immunol.*)

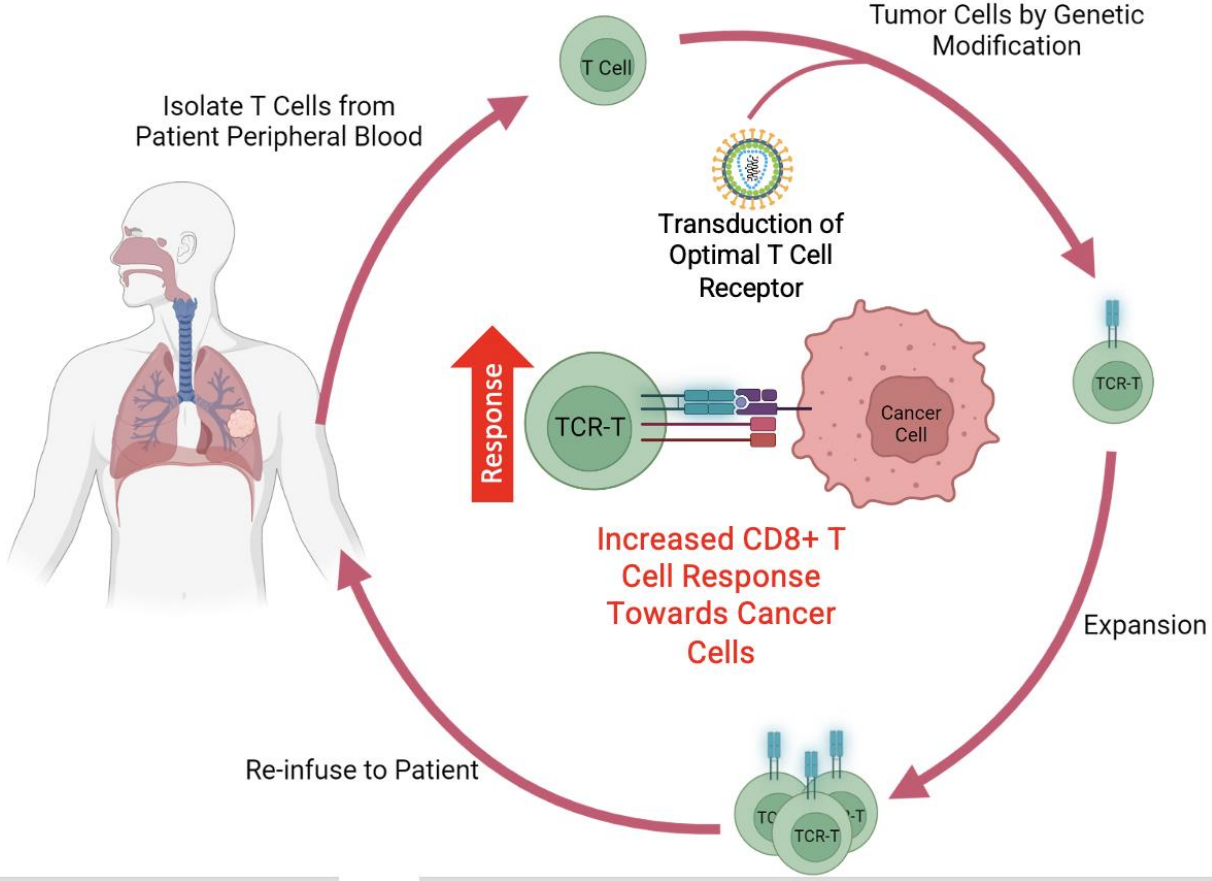
Castillo et al. (Unpublished)


Junction Crossing
Read 1
Read 2
Discordant
Gene 1
Gene 2
5' KIF5B [Exon 1-15]
RET [Exon 12-19] 3'
.....TGGGAAATAATGATGTAAAG
GAGGATCCAAAGTGGGAATT.....
K EDPKWEFP
VK EDPKWEF
DVK EDPKWE
NDVK EDPKW
NNDVK EDPK
GNNDVK EDP
VGNNDVK ED
AVGNNDVK E
Junction Crossing Neoantigenic Peptides


Personalized Vaccination
VKEDPKWEF
DVKEDPKWE
NDVKEDPKW
GNNDVKEDP
VGNNDVKED
AVGNNDVKE
KEDPKWEFP
NNDVKEDPK


Characterize Immune Response
Define Immune Populations
Activated Immune Cells
UMAP2
6
4
2
0
UMAP2
Active
Exhausted
Naive


UMAP1
Profile T Cell Receptor (TCR) Landscape
TCR Selection
Cell Fraction
0.008
0.006
0.004
0.002
0.000
Active
Exhausted
Naive
407
7939
394
5014
4210
5530
3978
4072
3968
142
4552
4708
7238
390
5130
Clonotype
Antigen Presenting Cancer Cell


Redirect T Cells to Target Tumor Cells by Genetic Modification
T Cell
Isolate T Cells from Patient Peripheral Blood
Transduction of Optimal T Cell Receptor
Response
TCR-T
Cancer Cell
TCR-T
Increased CD8+ T Cell Response Towards Cancer Cells
Expansion
Re-infuse to Patient
TCR-T

Breast Cancer Prevention 4-Site Sequencing

75 Patients

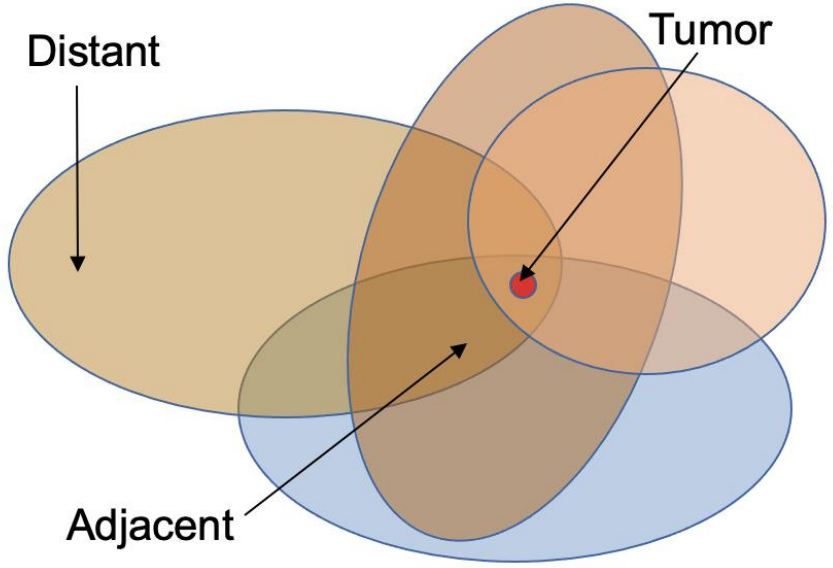
- 25 HR+
- 25 HER2+
- 25 TNBC

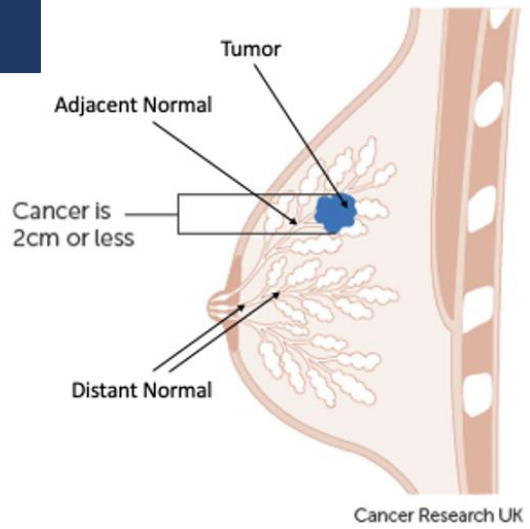
4 Sites

- Tumor
- Adjacent Normal
- 2 Distant Normal

4 Normal Controls

- ~25M read/sample





Phase I Study of a Personalized Breast Cancer Prevention Vaccine

Woman with precancer/at risk for cancer
- Atypia
- Lobular neoplasia

→ Sequencing of breast tissue → Bioinformatics analysis to identify the gene targets → Manufacture the vaccine against these gene targets → Vaccinate → Test whether vaccine generated immune response

Dr. Isabelle Bedrosian, MDACC

Fusion Discovery Pipeline

1. 75 Tumor samples from Sporadic 360sample cohort.
2. 43 BRCA1_2 mutated samples

Filtered for Driver Gene(COSMIC database) associated Gene Fusions only

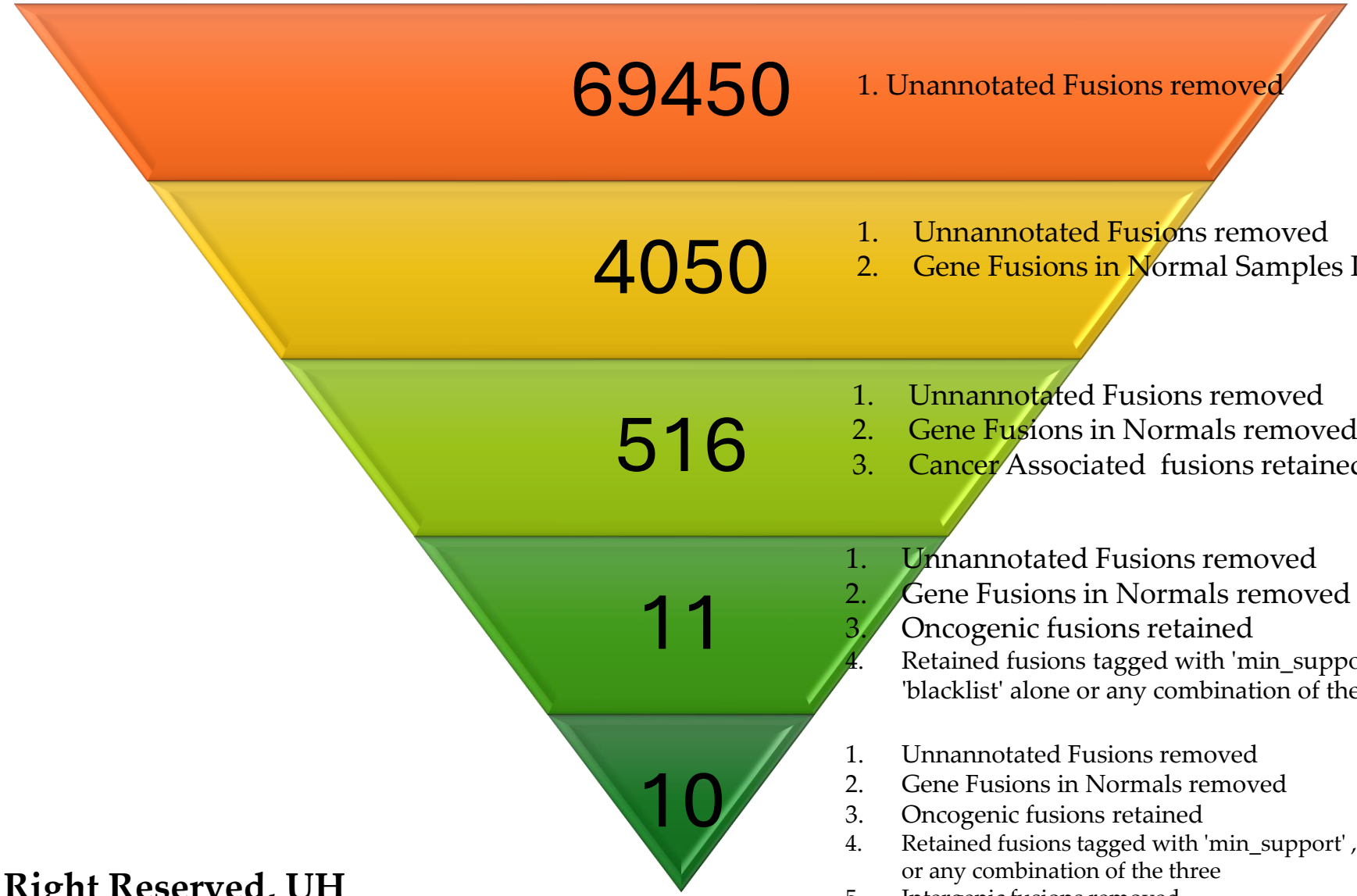
Removed Fusion Genes detected in 21 Normal Samples

Overlapping Fusion Genes were determined between the two datasets(shown above)

Overlapping Fusion Genes across subtypes were determined between the two datasets(shown above)

Fusions per sample were determined

Overlapping fusions across 75 Tumors and BRCA1_2 High Risk Dataset



mRNA Vaccine for Treatment & Prevention of Cancer

5' Cap	•Anti-Reverse Cap Analog (ARCA)
5' Untranslated Region	•High levels of expression for this UTR in humans.
Kozak Sequence	•Nucleic acid motif that functions as the protein translation initiation site in most eukaryotic mRNA transcripts.
Cassette	•Fusion Sequence
Stop Codons	•Standard Human Stop Codon
3' Untranslated Region	•High levels of expression for this UTR in humans.
Poly-A Sequence	•Polyadenylation signal sequence near end of 3' UTR

Peptide and mRNA Vaccine Development Team

Sakuni Rankothgedera

Lead Scientist
Single Cell and Spatial Transcriptomics Platform

CD74- NRG1

Lung Cancer
Ovarian Cancer

Aaranyah Kandasamy

Lead Scientist
PDX -Actionable Mutation Discovery Platform

ESR1-PRKN

Breast Cancer

Shiyanth Thevasagayampillai

Lead Scientist
mRNA Vaccine Discovery Platform

FGFR3-TACC3

Brain Cancer
Lung Cancer
Breast Cancer
Pancreatic Cancer
Ovarian Cancer

Dilshan Adikari

NGS Specialist
RNA & DNA Sequencing Platform

FGFR3-KHSRP

Pancreatic Cancer
Brain Cancer
Ovarian Cancer
Urothelial Cancer

Cole Woody

***In silico* fusion validation platform**

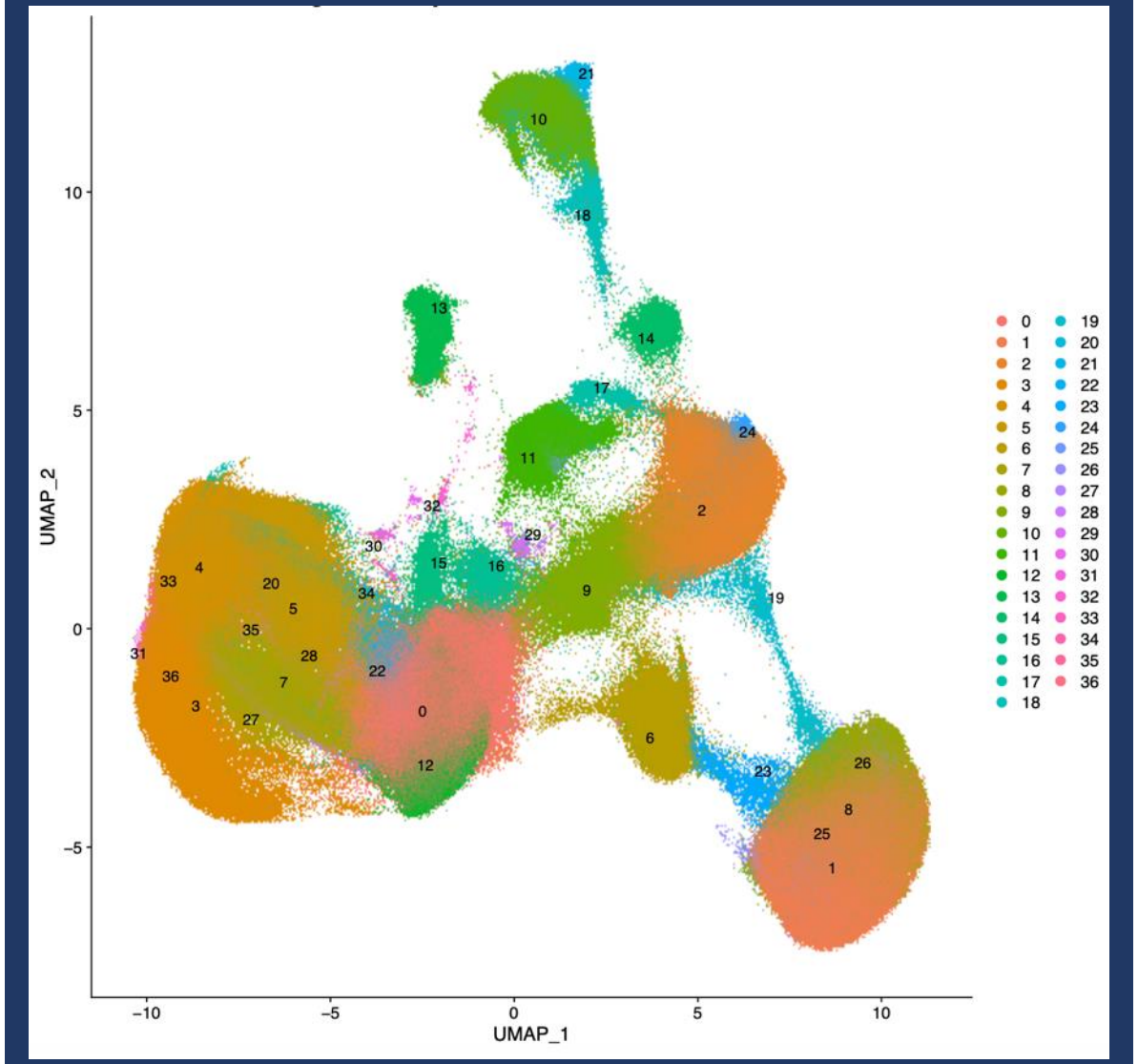
HNRNPM-CCNE1

Ovarian Cancer

UNIVERSITY of HOUSTON
College of Natural Sciences and Mathematics

UHSEQ Sequencing Core

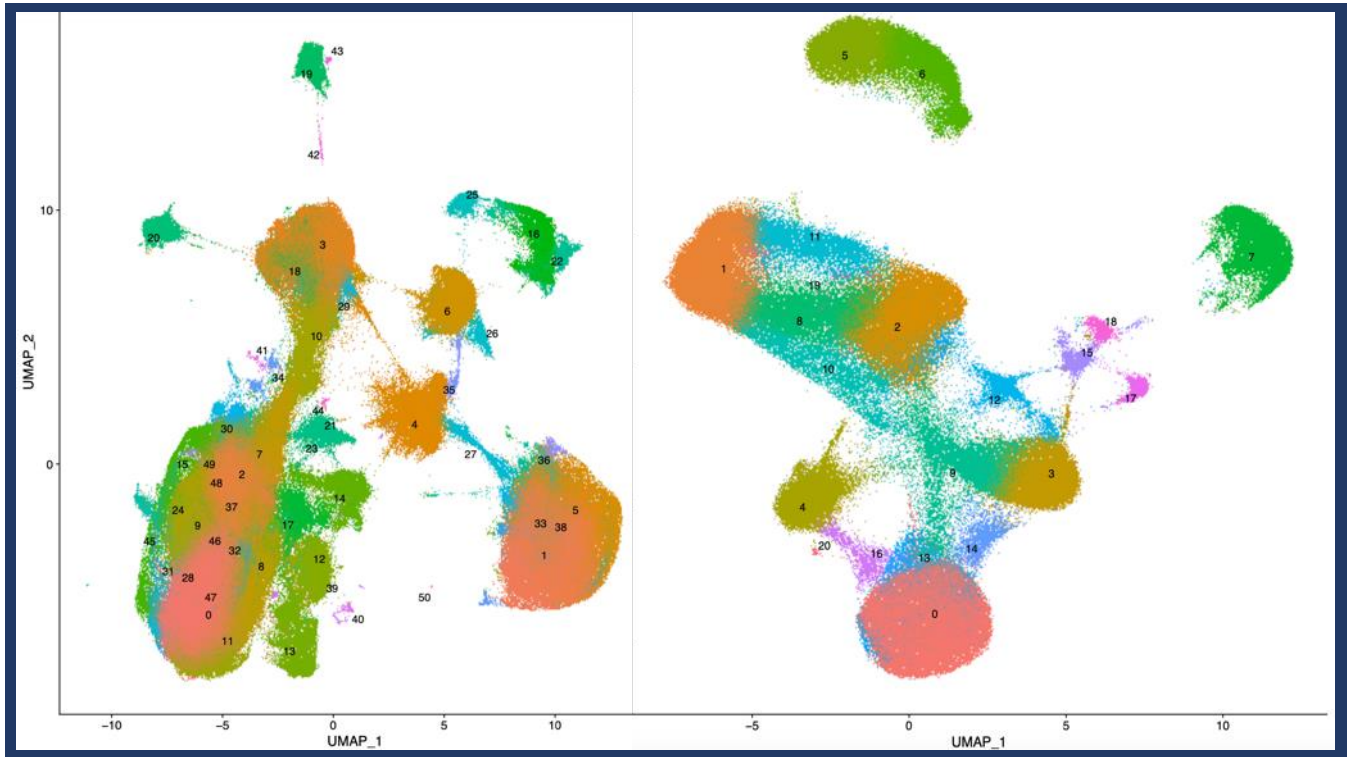
Sequencing a Million Cells from over a Hundred Brains from Opioid Use Disorder (OUD) Patients



1 Million Cells Profiled From Cryopreserved Brain

Hypothalamus	Thalamus	Globus Pallidus
~400,000 Cells	~300,000 Cells	~300,000 Cells

Hypothalamus | **Thalamus**

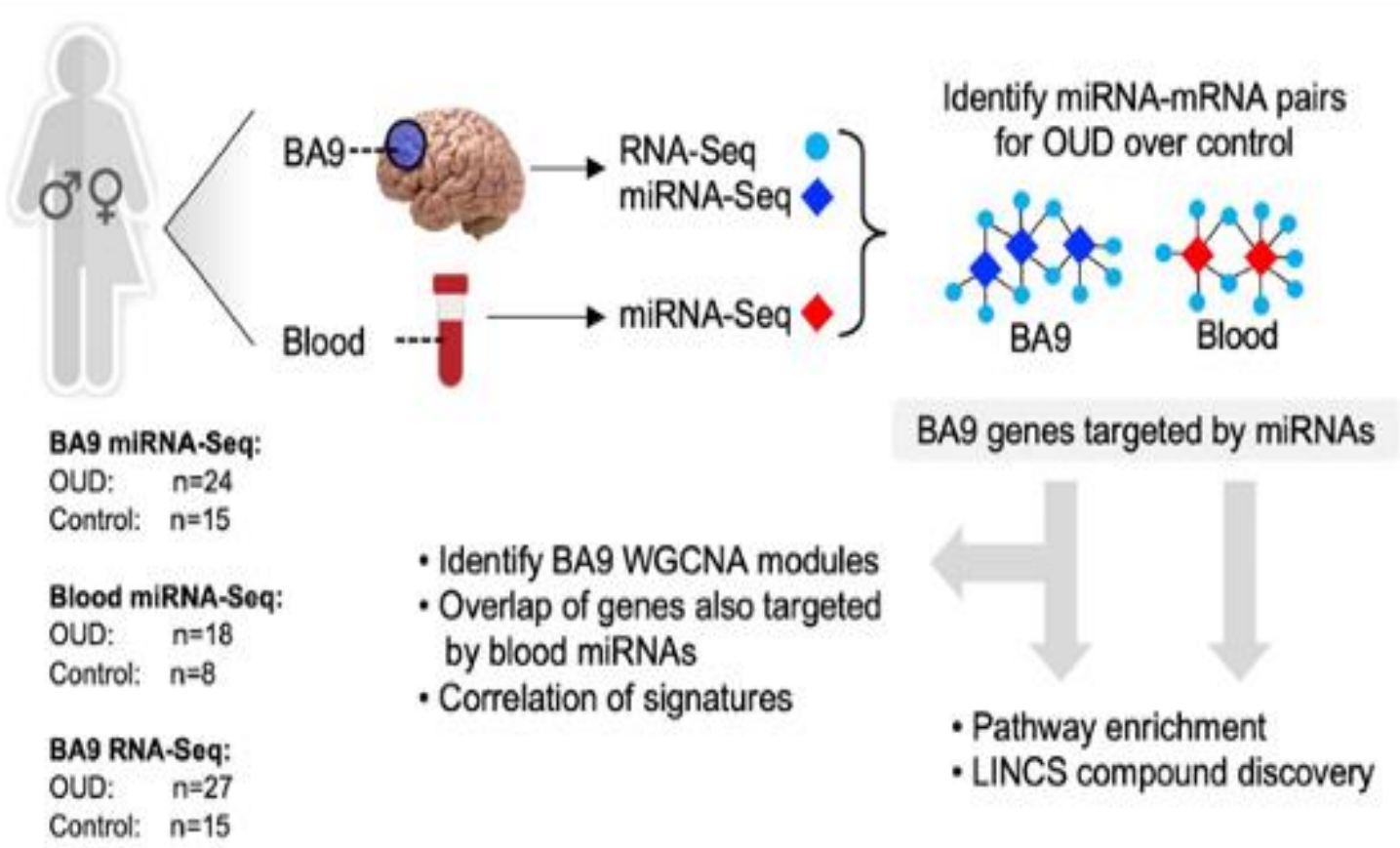


microRNA Biomarkers for Opioid Addiction Risk

Thor Loveland
Founder and CEO, miRMIND

miRMIND

- **miR-Tracker**
  - Blood Biomarker
  - Predict & prevent death from opioid overdose



UH Sequencing and Gene Editing Core

UHSEQ Sequencing Core

Administration

- Dr. Preethi Gunaratne –Director
- Sakuni Rankothgedera – Lead Scientist
- Aaranyah Kandasamy – Project Manager
- Shiyanth Thevasagayampillai – Manager Analytics Platforms

Business Office Liaisons

- Thien-Tam Do
- Malinthika Fernando
- Fred McGhee
- Judy Mata
- Rachel Marks

Core Collaborators

- **MD Anderson Cancer Center**
  1. Dr. Hussein Abbas
  2. Dr. Moran Amit
- **UT Health Science Center**
  1. Dr. Consuelo Walss-Bass – UT Health
- **Baylor College of Medicine**
  1. Dr. Andrew Dinardo
  2. Dr. Cristian Coarfa
- **University of Houston**
  1. Dr. Beau Alward – UH
  2. Dr. Mingfu Wu – UH
  3. Dr. Ashok Kumar – UH
  4. Dr. Sang-Hyuk Chung – UH
  5. Dr. Tasneem Bawa – UH
  6. Dr. Robert Schwartz – UH
  7. Dr. Meghana Trivedi – UH
- **Columbia University, NY**
  1. Dr. Emily Mace
- **Augusta University, GA**
  1. Dr. Huabo Su – Augusta University
  2. Dr. Hong Shi – Augusta University
  3. Dr. Xiaochun Long – Augusta University
- **University of South Florida, FL**
  1. Dr. Wanling Xuan – USF
- **Texas Tech University**
  1. Dr. Naima Moustaid-Moussa – TTU
- **Texas Childrens Hospital**
  1. Dr. Gabriel Loor
  2. Dr. Lourdes Chacon