Cancer Immunotherapy & Personalized Vaccines

Alex Rubinsteyn
Mount Sinai, Hammer Lab

February 21st, 2017

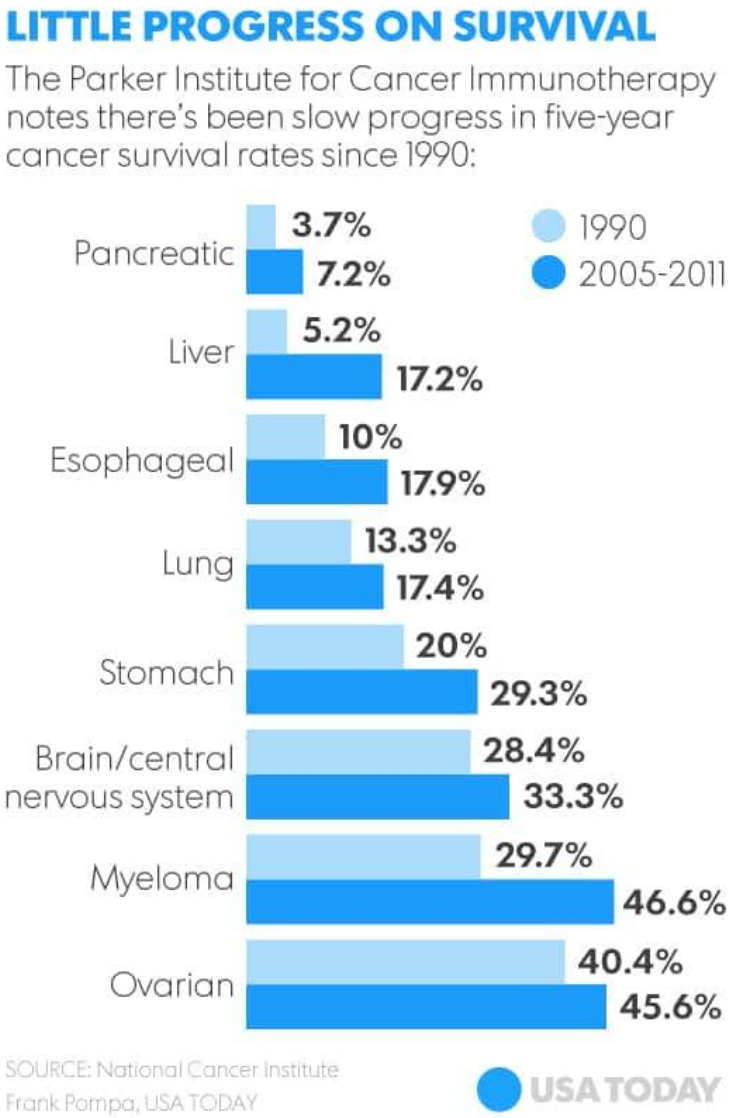
Hammer Lab @ Mount Sinai

- Backgrounds in math, compsci, ML, compbio
- Focus: *cancer immunotherapy*
  - Cancer genomics
  - Machine learning for immunology
  - Clinical trial analysis
- github.com/hammerlab

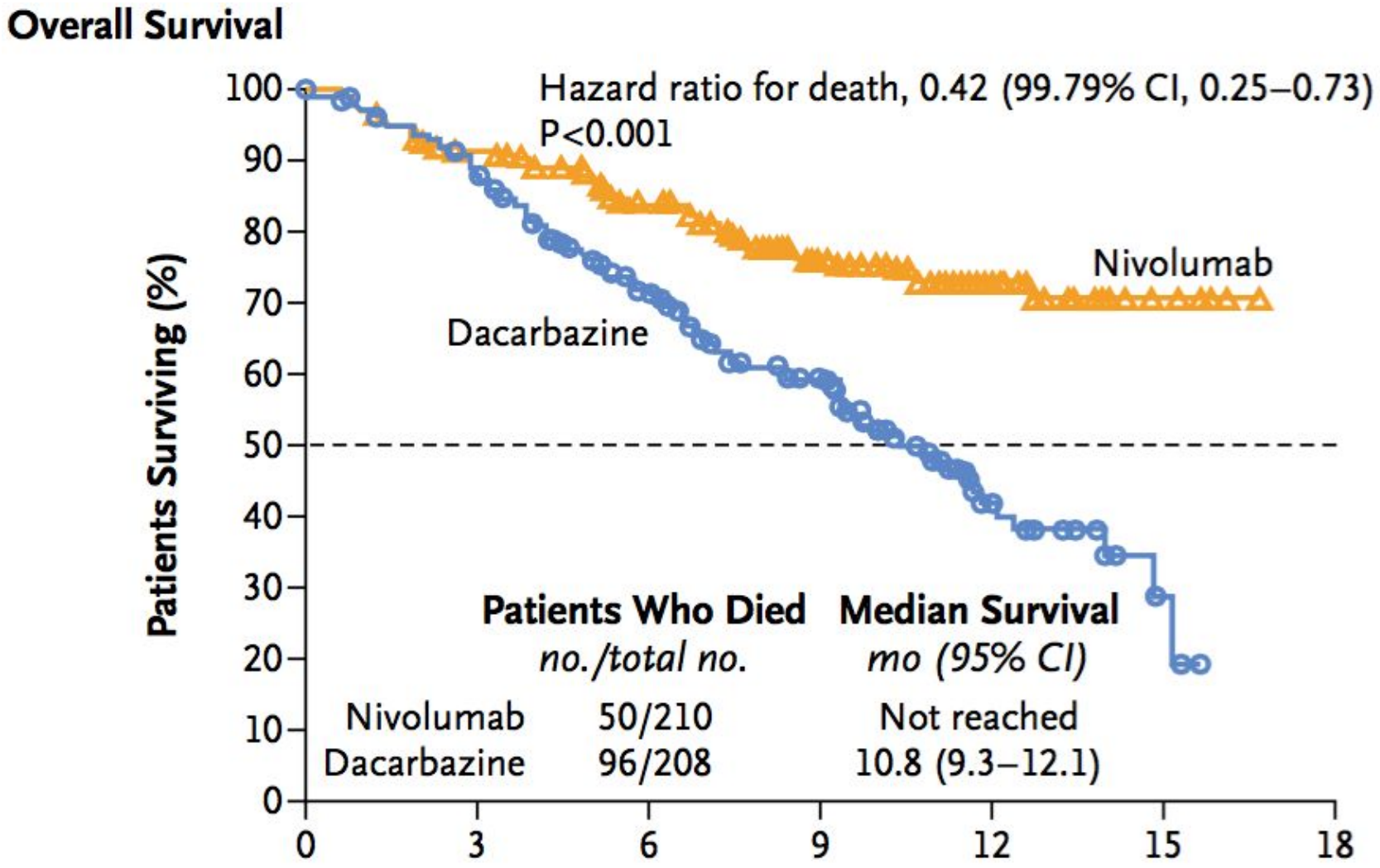
Icahn School of Medicine at Mount Sinai

PARKER INSTITUTE for CANCER IMMUNOTHERAPY

Great PFS improvements but…



Cancer Immunotherapy



Nivolumab in Previously Untreated Melanoma… (Robert NEJM 2015)

Flavors of Cancer Immunotherapy

Checkpoint blockade	Cellular therapies	Vaccines
Disinhibit CD8+ T-cells, antigens responsible for tumor clearance unknown. <u>Success stories:</u> • αCTLA-4 (ipi) • αPD-1 (pembro, nivo) • αPD-L1 (atezo)	Ex-vivo expansion of patient T-cells after receptor engineering and/or selection. <u>Success stories:</u> • CD19 CAR T-cells for B-cell malignancies	<u>Success stories:</u> • ??? Tumor-associated vaccines haven't panned out but "neoantigen" vaccines are a hot field. <u>Contenders</u>: Cathy Wu @ DFCI, Nina B @ MSH, Ugur Sahin @ TRON, Neon, Gritstone, Caperna, &c

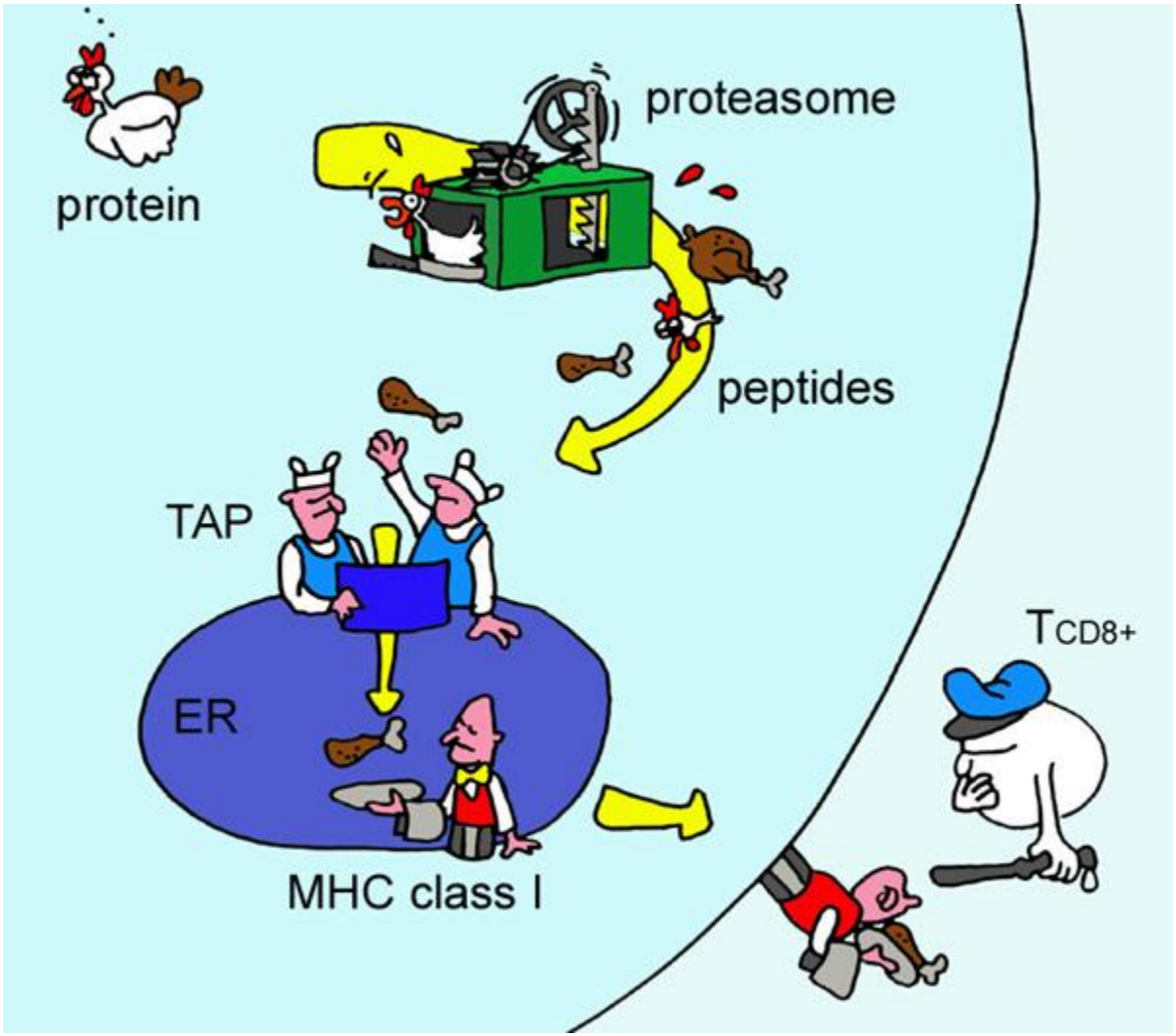
Ingredients of a Neoantigen Vaccine

- **Neoantigens**
  - Often mutant peptides presented on Class I MHCs of tumor cells
  - Fancy: larger indels, structural variants, chromosomal fusions
- **Vehicle**
  - Peptides
  - Dendritic cells pulsed with peptides
  - mRNA (bare or in a liposome)
  - DNA (?)
- **Adjuvant**
  - TLR agonist (e.g. poly-ICLC)
  - Oil emulsion (e.g. montanide)

Identifying Neoantigens

- **Somatic mutations**
  - DNA sequencing of tumor sample & normal (PBMCs)
  - Usually SNVs & small indels (limited by read length, aligner, variant caller)
  - Fancy: larger indels, structural variants, chromosomal fusions
- **Evidence that mutation is expressed**
  - RNA sequencing of tumor sample
  - Bewilderingly omitted from many neoantigen analyses
- **Immunological predictions**
  - Binding of mutant peptides to patient's Class I MHCs (i.e. NetMHC)
  - Antigen processing: proteasomal cleavage, transport into endoplasmic reticulum, trimming by ER peptidases
  - Class II MHC pathway: full of unknown unknowns
- **(optional) Elution of peptides + mass spectrometry**
  - Assay bias (e.g. can't detect Cysteines), sensitivity unclear

Antigen Processing in 5 seconds



Source: Morten Nielsen

Class I MHC Binding Prediction

- **MHC Genes**
  - Highly variable locus with 1000s of alleles in the human population
  - Classical genes for presenting peptides to CD8+ T-cells: HLA-{A, B, C}
  - “HLA typing” means determining the 2x HLA-A, 2x HLA-B and 2x HLA-C alleles each person carries
- **Models**
  - Input: peptide sequence
  - Output: scalar value (affinity or stability)
  - Commonly used: neural network affinity predictors (NetMHC, NetMHCpan)
- **Data Source**
  - Immune Epitope Database (iedb.org)
  - Curate a variety of in-vitro peptide:MHC affinity and stability assays

PGV001: Safety and Immunogenicity of Personalized Genomic Vaccine to Treat Solid Tumors

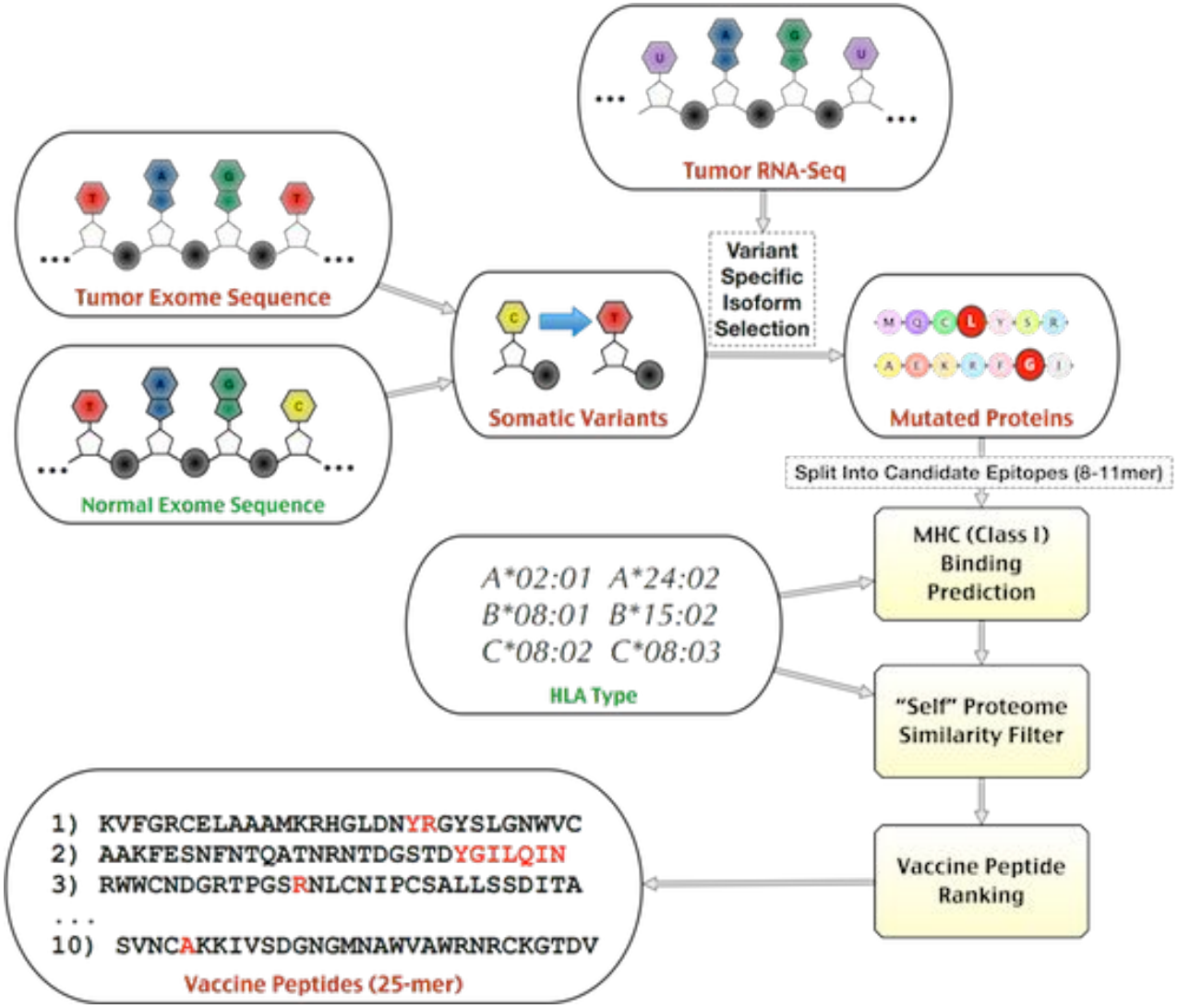
(Phase I Clinical Trial at Mount Sinai)

Nina Bhardwaj

Personalized Genomic Vaccine

- Resectable solid tumors without evidence of metastatic disease
  - H&N, NSCLC, Breast, Ovarian, Urothelial, SCC
- 10x peptides (~25 amino acids)
- Expressed somatic mutations
- Predicted to generate Class I MHC ligands
- Adjuvant: Poly-ICLC
  - TLR3 agonist
  - Similar to dsRNA but inosine:cytidine polymer (instead of guanosine), stabilized with lysine
- Endpoint: safety and feasibility

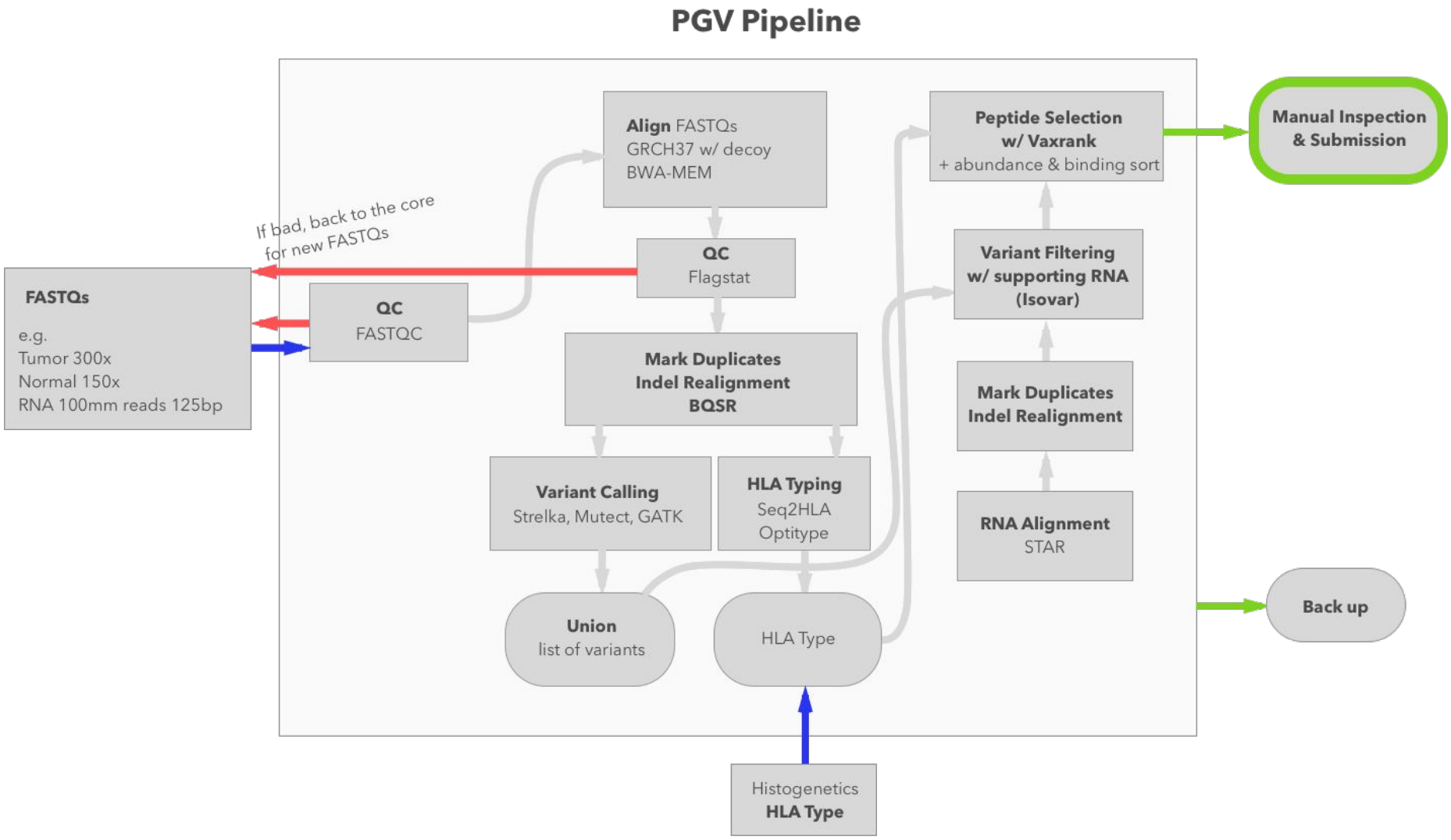
Tumor Neoepitope Selection



Sequencing Details

- Prefer fresh frozen samples
  - Formalin and/or age both conspire to make your sample useless
- Exome capture kit: SureSelect XT
  - Tagmentation WES kits (e.g. QXT) create many duplicate reads
- Normal DNA: extracted from blood, 150x mean coverage
  - High mean coverage necessary due to non-uniformity of capture kit
- Tumor DNA: 300x tumor mean coverage
  - Expecting ~50% sample purity, 150x * 2 = 300x
- Tumor RNA: ~100-150M mRNA reads
  - Older FFPE samples give us <20% reads mapping to annotated exons

Pipeline


PGV Pipeline
FASTQs
e.g.
Tumor 300x
Normal 150x
RNA 100mm reads 125bp
If bad, back to the core for new FASTQs
QC
FASTQC
Align FASTQs
GRCH37 w/ decoy
BWA-MEM
QC
Flagstat
Mark Duplicates
Indel Realignment
BQSR
Variant Calling
Strelka, Mutect, GATK
HLA Typing
Seq2HLA
Optitype
Union
list of variants
HLA Type
Histogenetics
HLA Type
RNA Alignment
STAR
Mark Duplicates
Indel Realignment
Variant Filtering
w/ supporting RNA
(Isovar)
Peptide Selection
w/ Vaxrank
+ abundance & binding sort
Manual Inspection
& Submission
Back up

Tools Developed for the Trial

varcode	Python interface for VCFs, variant effect prediction
isovar	Determine mutant coding sequence from RNA-seq
vaxrank	Vaccine peptide selection (including manufacturability)
epidisco	Turn-key workflow to generate vaccine peptide report from FASTQ inputs (runs all bioinformatics tools)
ketrew	Workflow engine used to run tools on Google Cloud
mhctools	Standard interface to pMHC binding predictors
pyensembl	Python interface to Ensembl reference genome annotations

Available at github.com/hammerlab

Coding Sequence Prediction (varcode)

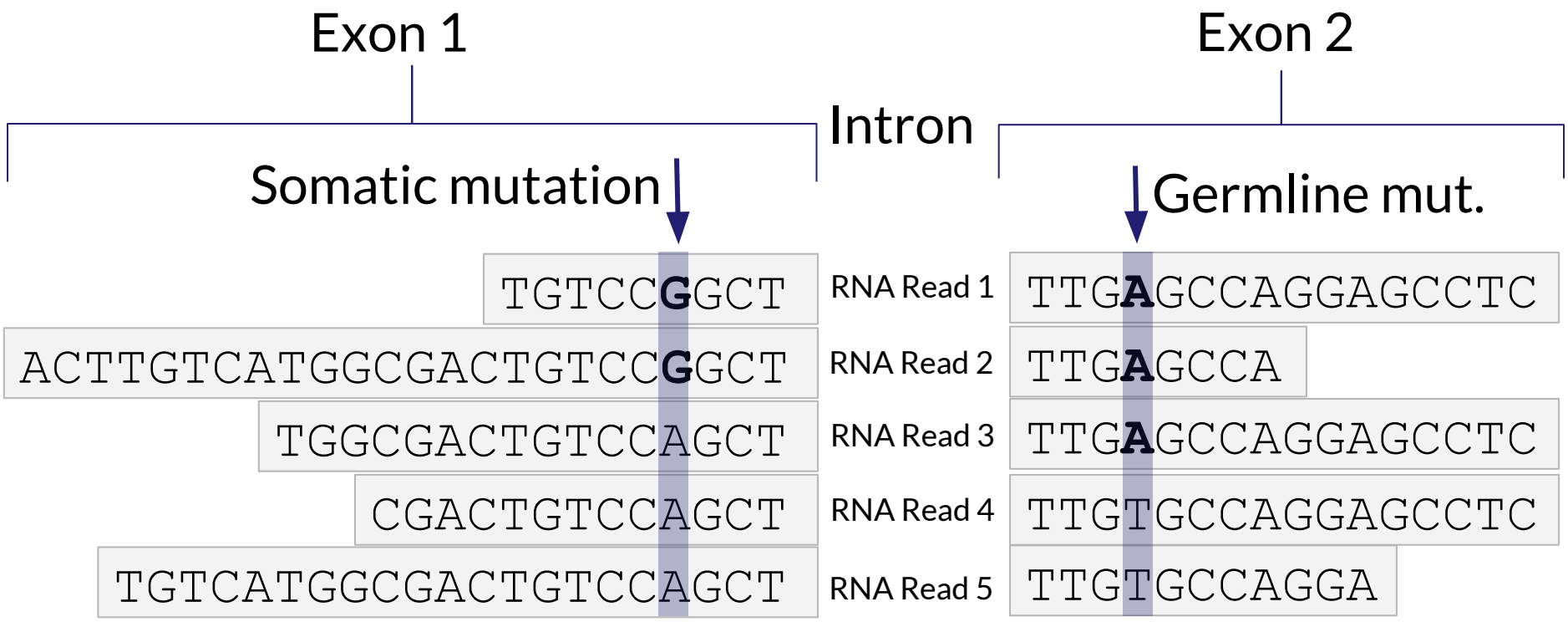
Code	Value

`variant = varcode.Variant(`
`  "3",`
`  36779850,`
`  ref="C",`
`  alt="",`
`  ensembl='grch37')`
	`Variant(`
	`  contig='3',`
	`  start=36779850,`
	`  ref='C',`
	`  alt='',`
	`  reference_name='GRCh37')`

`effect =`
`  variant`
`  .effects()`
`  .top_priority_effect()`
	`FrameShift(`
	`  variant=chr3 g.36779850_36779850delC,`
	`  transcript_name=DCLK3-001,`
	`  transcript_id=ENST00000416516,`
	`  effect_description=p.E101fs)`

`effect.mutant_protein_sequence`	MGKEPLTLKSIQVAVEELYPNKARALTLAQHSRAPSPRLR SRLFSKALKGDHRCGETETPKSCSEVAGCKAAMRHQGKIP EELSLDDRARTQKKWGRGKW **SQNPVASPPGKPLWKRGTQG ERSILGWRLKRPRVKLSDARSARERGSSSRAWSVRGFLWG PVSWIWGRAQCMMWRSW**

Phasing and Transcript Selection (isovar)



Selected coding sequence includes germline mutation:

GGCGACTGTCCGGCTTTGAGCCAGGTGCCTC

Vaccine Peptide Selection (vaxrank)

```
vaxrank
   --vcf mutect.vcf
   --vcf strelka.vcf
   --bam tumor-rna.bam
   --vaccine-peptide-length 25
   --mhc-predictor netmhcpan
   --mhc-alleles-file
   alleles.txt
```

1.

Variant	chr12 g.82752062C>T
Gene name	CCDC59
Top score	38.3
Reads supporting variant allele	32
Reads supporting reference allele	152
Reads supporting other alleles	0

Predicted Effect

Effect type	Substitution
Transcript name	CCDC59-001
Transcript ID	ENST00000256151
Effect description	p.V32M

Vaccine Peptides for variant: chr12 g.82752062C>T

i.

Amino acid sequence	STVGYRNKN M RQKTWRPNHPQAFVG		
Transcript name	CCDC59-001		
Length	25		
Expression score	5.66		
Mutant epitope score	6.77		
Reads fully spanning cDNA sequence(s)	2		
Mutant amino acids	1		
Mutation distance from edge	9		
	Sequence	IC50	Allele
	KNMRQKTWR	27.95	HLA-A*31:01
	NMRQKTWR	36.99	HLA-A*31:01
	RNKNMRQKTWR	63.01	HLA-A*31:01
	STVGYRNKNMR	95.81	HLA-A*31:01

Sufficient MHC expression?

	Dry Run Patient #1	Dry Run Patient #2	Dry Run Patient #3	Dry Run Patient #4	Dry Run Patient #5
HLA-A	33.6 (FPKM)	1.1	60.7	74.1	41.2
HLA-B	46.5	0.7	136.9	59.5	63.7
HLA-C	65.8	0.7	97.7	55.5	89.1

Do we get enough mutations?

	Dry Run Patient #1	Dry Run Patient #2	Dry Run Patient #3	Dry Run Patient #4	Dry Run Patient #5
Total Variants	501	888	591	663	912
(Non-silent) Coding Variants	180	253	173	231	305
Frame Shifts	4	8	1	3	1
Total Peptides in Report	11	9	17	32	22
Peptides with Predicted MHC ligands of affinity <= 100nM	4	3	8	10	9

Can the peptides be manufactured?

Solid-phase peptide synthesis (and purification) fail for many peptides, try to anticipate with these scores:

```
# Priority I: GRAVY score of 7 residues closest to the C terminus
cterm_7mer_gravy_score,

# Priority II: GRAVY score of any 7mer window in the peptide sequence
max_7mer_gravy_score,

# Priority III: avoid N-terminal Gln, Glu, Cys
difficult_n_terminal_residue,

# Priority IV: avoid C-terminal Cys
c_terminal_cysteine,

# Priority V: avoid C-terminal Pro
c_terminal_proline,

# Priority VI: total number of Cys residues
cysteine_count,

# Priority VII: avoid N-terminal Asn
n_terminal_asparagine,

# Priority VIII: avoid Asp-Pro bonds
asparagine_proline_bond_count,
```

First Patient Timeline

- Oct 5, 2016 - samples acquired
- Oct 10 - pathology deposits samples in Genomics Core
- Oct 17 - sequencing data delivered
- Oct 19 - vaccine pipeline completes
  - 9 credible neoepitope-generating non-synonymous mutations identified
- Oct 20 - Histogenetics (HLA types) report arrives
  - Concordant with seq2hla except for one HLA-C allele

….many bioinformatics tweaks & fixes later…

- December: 14 usable variants!
- January: Still talking with manufacturer about whether we can use peptides containing Cysteines (oxidation!)

PGV001 Patients

- PT001 - withdrew due to recurrence
- PT002 - H&N
  - will be vaccinated in March
- PT003 - Ovarian clear cell carcinoma
  - Waiting for FDA approval of IND amendment
- PT004 - H&N
  - Crude peptide synthesis starting soon
- PT005 - ER-/PR-/Her2- breast cancer
  - Just finished sequencing tumor + normal samples

End