SNAQ™-SEQ INTERNAL STANDARDS: QUALITY CONTROL TECHNOLOGY THAT IMPROVES NGS ACCURACY

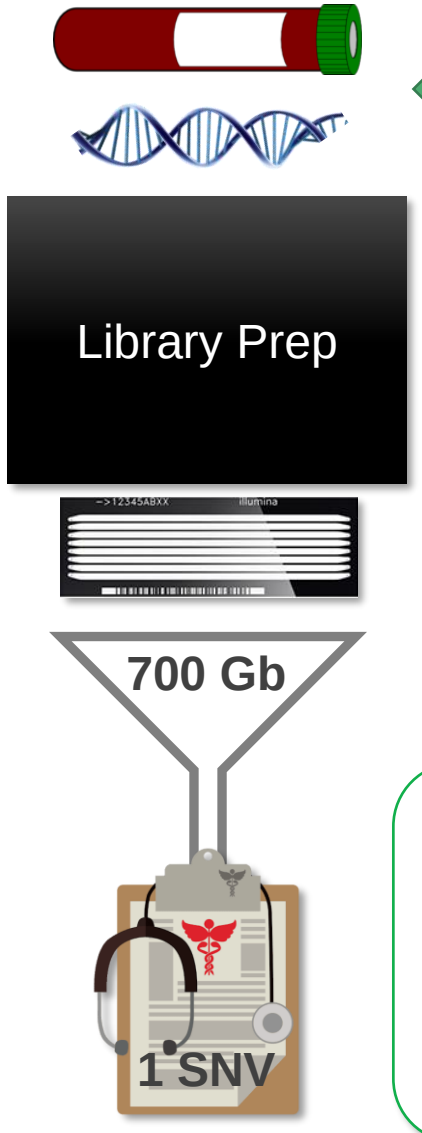
AccuGenomics

A higher standard of accuracy

www.accugenomics.com

NGS IS A CRITICAL BUT COMPLEX TOOL

- Flexible tool for detecting clinically actionable variants
- Overkill, 0.07% sequence relevance
- Complex testing procedure
  - Microscale fluidics & detection
  - Bioinformatics on Gb sequence
- Modest pass/fail QC procedures
  - Sample input level
  - Insert yields
  - Flowcell metrics
  - PhiX & reference sample
  - Unique Molecular Indices (UMI)



SNAQ-SEQ Internal Standards

140 CONSUMABLES

362 STEPS

52 INSTRUMENT INTERACTIONS

TRILLIONS OF REACTIONS

MANY BIOINFORMATIC STEPS

- **Validation is *not adequate assurance* of tomorrow's performance**
- **Good QC critical for low signal events: ctDNA, exon loss, limits of detection**

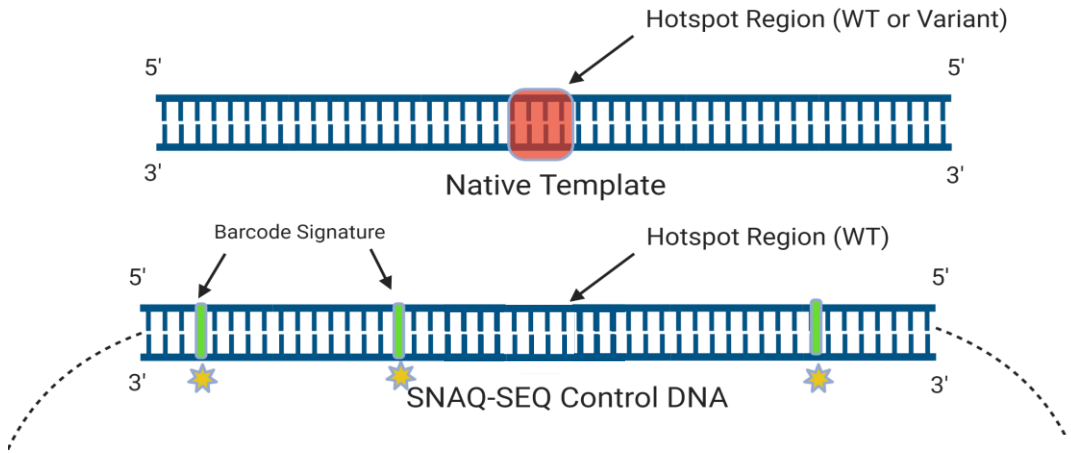
AccuGenomics

SNAQ-SEQ
LIMIT OF BLANK

- ✓ QC for Variant Calling
- ✓ Increased Accuracy

AccuGenomics

STANDARDIZED **N**UCLEIC **A**CID **Q**UANTIFICATION FOR **SEQ**UENCING (SNAQ-SEQ)



- Based on use of **I**nternal **S**tandards (IS)
  - Designed to clinical ROI
  - Reference sequence manufactured with 10^{-8} error rate
  - Intermittent base changes enable bioinformatic separation
  - Biochemically mimic sample except issues arising from pre-damaged DNA (e.g., FFPE)
  - Added to every sample prior to library prep

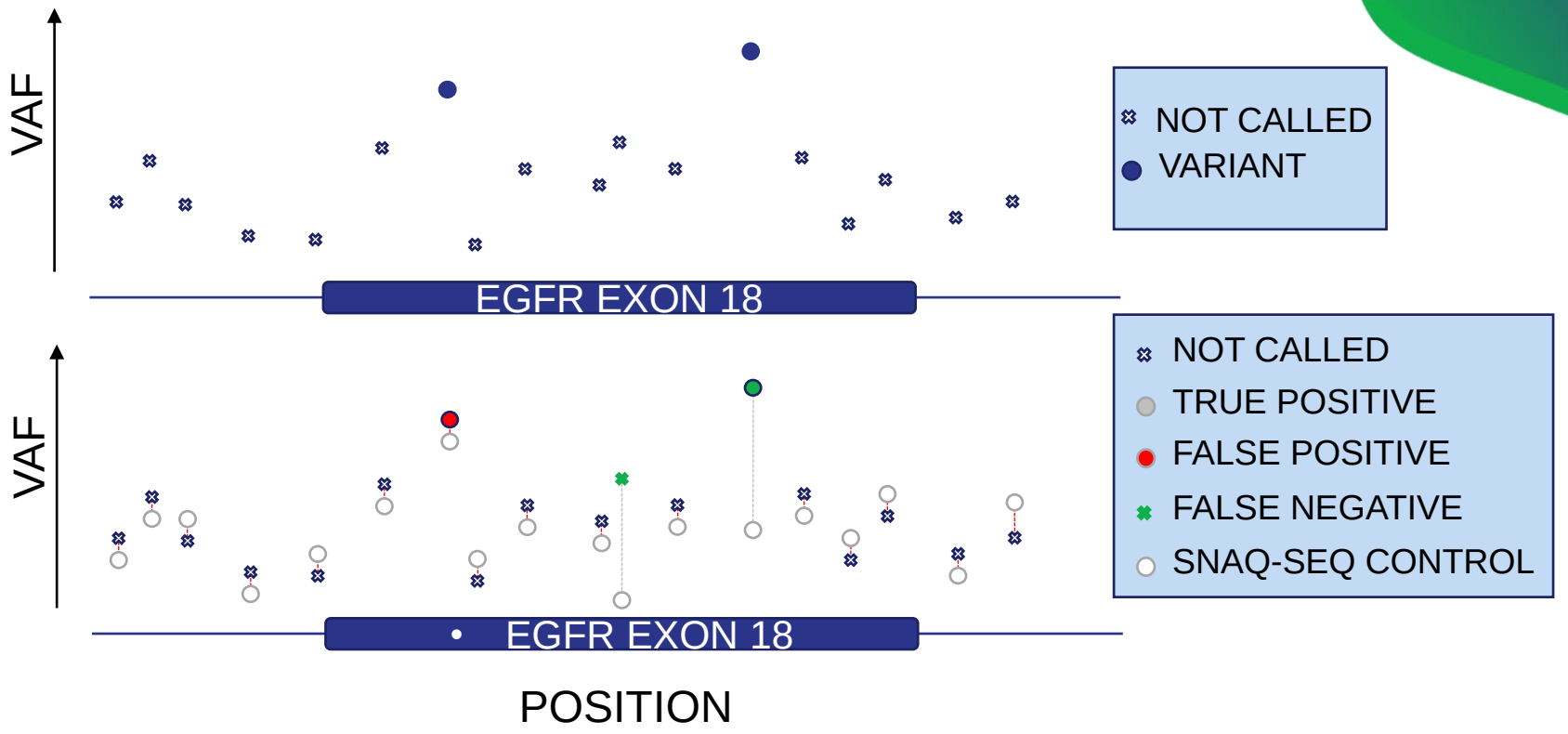
- Limit of Blank
  - Ultimate negative control
  - Mimic sample sequencing errors
  - Applications: low VAF
- Accurate quantification
  - Ratio between sample and IS maintained
  - Knowing IS input and ratio enables accurate quantification of input template
  - Applications: CNV, ctDNA/ml plasma, TA

Quality Control for EVERY variant in EVERY sample

AccuGenomics

SNAQ-SEQ *LIMIT OF BLANK*: HOW IT WORKS

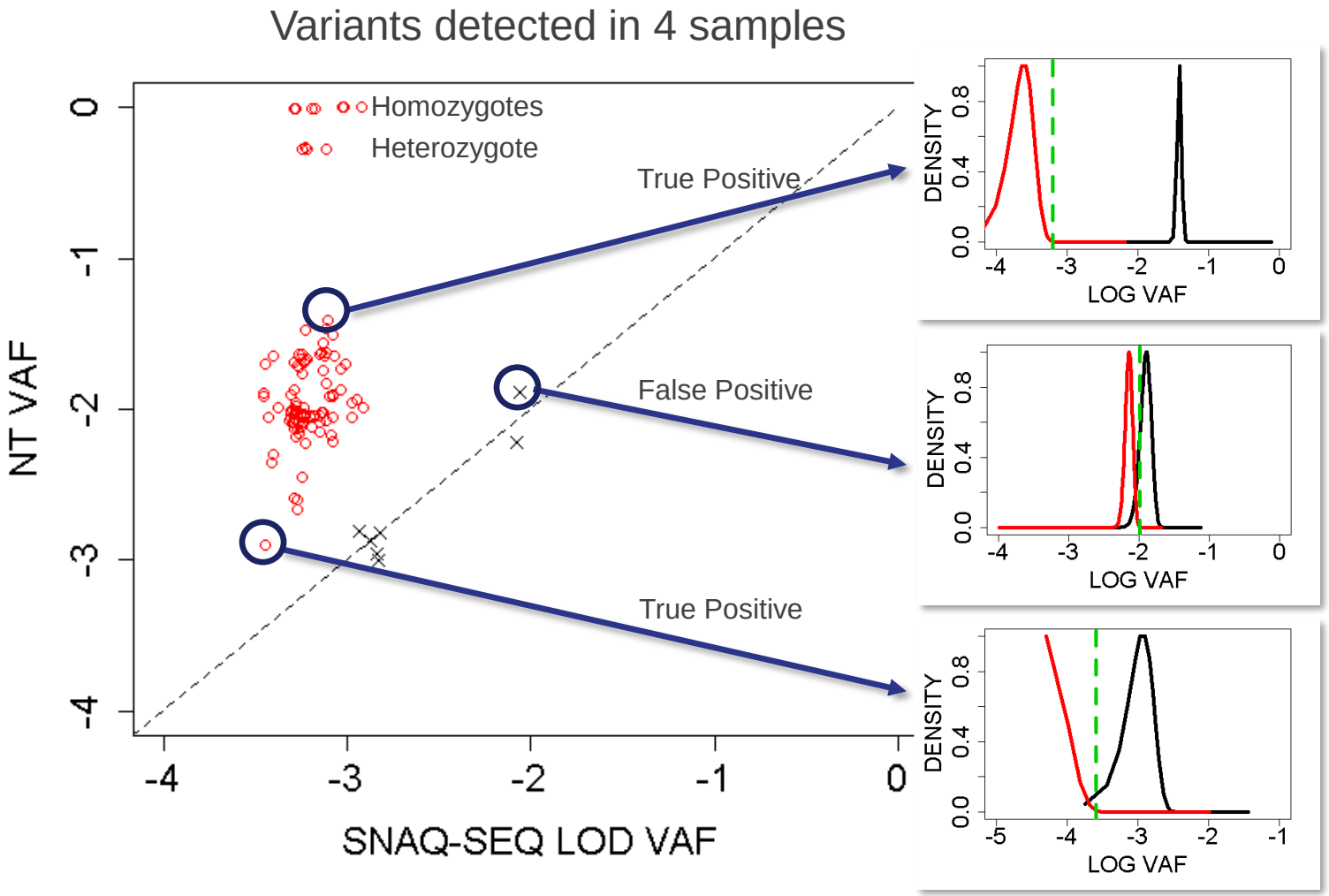
- Variant calling use probabilistic and heuristics methods
- Current low VAF approaches established during development
  - tumor/normal
  - Panel of normal
  - VAF cutoff
  - UMI
- Variant callers do not handle sample variability or technical method drift
- SNAQ-SEQ determines significant difference between sample and IS variant
- Any variant in the IS indicates a sequencing error



SNAQ-SEQ allows calculation of the significance above background error for *each* variant in *every* sample

AccuGenomics

SOMATIC VARIANT DETECTION IN ctDNA



- 25ng contrived ctDNA samples
  - 10 tumor cell line mixture
  - Diluted 5-fold normal genome
  - Fragmented to 150 bp
- Spiked with cfDNA IS
  - 32Kb flanking actionable mutations
  - Fragmented to 150 bp
- Illumina TST-170 library prep with UMI
  - 27 SNV covered by IS
- SNAQ-SEQ analysis of variants
  - Poisson Exact Test used to determine significance
  - Significance cutoff set by Bonferroni adjusted 5% alpha

SNAQ-SEQ eliminates false positive variants

"Assessing synthetic reference sequence internal standards as quality-control for NGS measurement of actionable mutations in circulating tumor DNA" in preparation SEQC2 Workgroup #2 using Accukit™ SEQC2MIX4 Catalog #1154

AccuGenomics

SNAQ-SEQ QC EXAMPLE REPORT

GENE	CHROM	POS	REF	OBS	AA_MUT	COSMIC_ID	1:5 ctDNA				1:25 ctDNA			
							1	2	3	4	1	2	3	4
NRAS	chr1	115256529	T	A	p.Q61L	COSM583	0.90%	0.89%	0.62%	0.68%	0.34%			
NRAS	chr1	115258748	C	A	p.G12C	COSM562	0.87%	1.03%	0.85%	0.92%				
NRAS	chr1	115258745	C	A	p.G13C	COSM570								
NRAS	chr1	115258745	C	G	p.G13R	COSM569								
NRAS	chr1	115258745	C	T	p.G13S	COSM571								
MAP2K1	chr15	66729162	C	T	p.P124S	COSM235614	1.02%	0.75%	0.88%	0.90%				
MAP2K1	chr15	66729136	T	C	p.L115P	NA								
TP53	chr17	7577085	C	T	p.E285K	COSM10722	0.45%	0.50%	0.59%	0.36%				
TP53	chr17	7577118	C	A	p.V274F	COSM10769	2.30%	2.05%	2.26%	2.72%	0.48%	0.50%		0.49%
TP53	chr17	7578211	C	T	p.R213Q	COSM10735	2.01%	1.91%	2.06%	1.96%	0.42%	0.30%	0.40%	0.36%
TP53	chr17	7577141	C	A	p.G266V	COSM10958								
TP53	chr17	7577141	C	T	p.G266E	COSM10867								
TP53	chr17	7577550	C	A	p.G244V	COSM43652								
TP53	chr17	7577550	C	T	p.G244D	COSM10883								
TP53	chr17	7578395	G	A	p.H179Y	COSM10768								
TP53	chr17	7578475	G	A	p.P152L	COSM10790								
PIK3CA	chr3	178936091	G	A	p.E545K	COSM763	1.21%	1.25%	1.03%	1.20%				
PIK3CA	chr3	178936091	G	C	p.E545Q	COSM27133								
PIK3CA	chr3	178936092	A	G	p.E545G	COSM764								
PIK3CA	chr3	178936092	A	C	p.E545A	COSM12458								
CTNNB1	chr3	41266125	C	T	p.T41I	NA								
PIK3CA	chr3	178936074	C	G	p.P539R	COSM759								
PIK3CA	chr3	178936082	G	A	p.E542K	COSM760								
PIK3CA	chr3	178936083	A	T	p.E542V	COSM762								
PIK3CA	chr3	178936093	G	T	p.E545D	COSM765								
PIK3CA	chr3	178936094	C	A	p.Q546K	COSM766								
PIK3CA	chr3	178936094	C	G	p.Q546E	COSM6147								
PIK3CA	chr3	178936095	A	G	p.Q546R	COSM12459								
PIK3CA	chr3	178936095	A	C	p.Q546P	COSM767								
EGFR	chr7	55259485	C	T	p.P848L	COSM22943								
MET	chr7	116412043	G	C	p.D1010H	COSM5574327								
MET	chr7	116412043	G	T	p.D1010Y	COSM3182								
FGFR3	chr4	1803568	C	G	p.S249C	COSM715								

- SEQC-2 Workgroup 2 contrived ctDNA samples
- Study concluded not to go below 0.5% VAF
- Report mocks up a 0.5% VAF requirement for a subset of 209 hotspot mutations covered by the IS

✓ Variants passed SNAQ-SEQ QC (green)

✓ Position LOB exceeds 0.5% VAF (orange)

✓ Insufficient coverage for 0.5% VAF (blue)

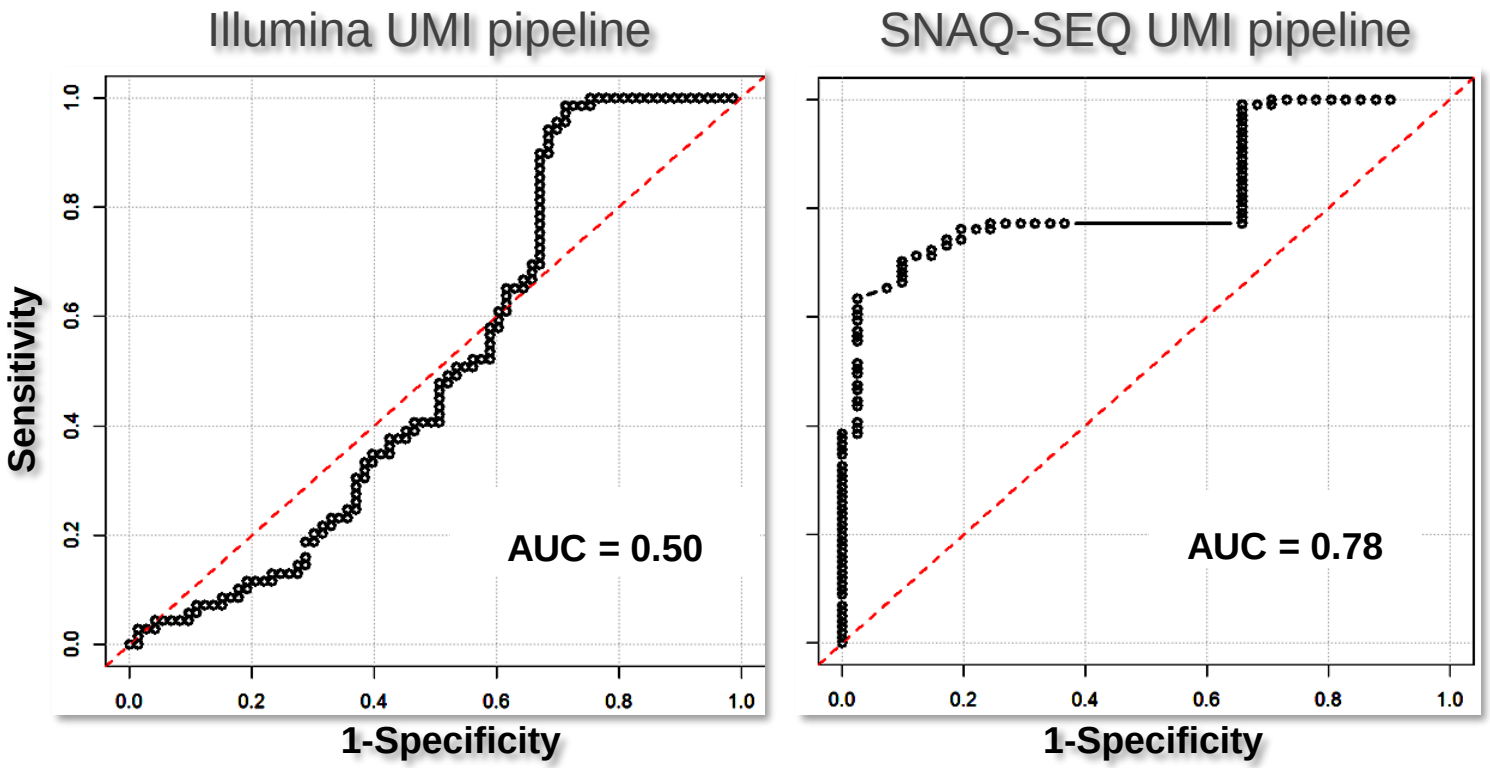
SNAQ-SEQ provides independent Quality Control for *each hotspot variant in every sample*

"Assessing synthetic reference sequence internal standards as quality-control for NGS measurement of actionable mutations in circulating tumor DNA" *in preparation* SEQC2 Workgroup #2 using Accukit™ SEQC2MIX4 Catalog #1154

AccuGenomics

SNAQ-SEQ IMPROVES ACCURACY

- cfDNA samples ranging 0.1% to 0.5% VAF
- Illumina TST-170 specificity varied by VAF cutoff
- SNAQ-SEQ specificity varied by Poisson Exact Test (PET) significance

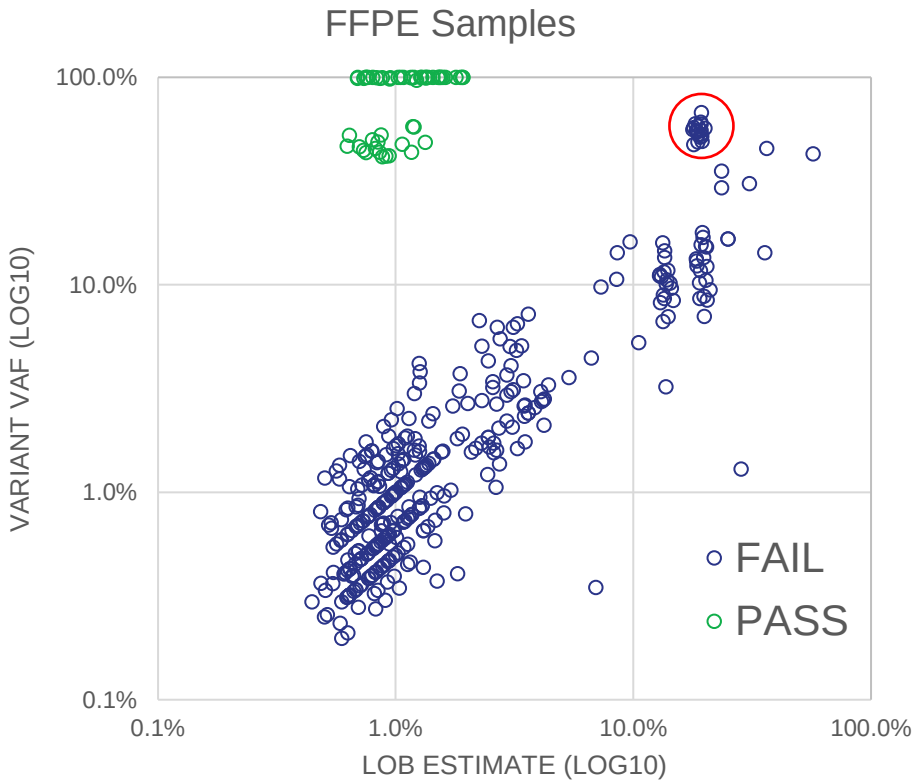


SNAQ-SEQ significance cutoff performs better than TST-170 VAF

"Assessing synthetic reference sequence internal standards as quality-control for NGS measurement of actionable mutations in circulating tumor DNA" *in preparation* SEQC2 Workgroup #2 using Accukit™ SEQC2MIX4 Catalog #1154

AccuGenomics

RESCUE OF POOR QUALITY FFPE SAMPLES



- ‘Normal’ FFPE samples were provided to SEQC2 NGS vendors as a simple sample to demonstrate false positive arising from FFPE
- ✓ DNA input was 10% expected which led to high false positive rates among all vendor platforms
  - *Samples were thrown out of study*
- SNAQ-SEQ IS spiked-in at very low levels
- PET cutoff set by eliminate IS false positives
- ✓ SNAQ-SEQ eliminated all false positives but not all variants were rescued due to suboptimal coverage

SNAQ-SEQ “rescues” low input FFPE samples with potential to give a result on ANY sample

Data from: “Identification of key quality control factors that affect targeted NGS variant calling of FFPE processed samples” *in preparation*
Study done with SECQC-2 Targeted Sequencing Workgroup using Accukit™ SEQC2Mix4 Catalog # 1154

AccuGenomics

WILL SNAQ-SEQ ENABLE TUMOR VARIANT ABUNDANCE PER PLASMA VOLUME MEASUREMENT?

- Plasma cfDNA varies 2-logs (5 to 1000 ng/ml)[1]
- VAF based monitoring affected by cfDNA levels
- Solution: measure variant/ml plasma using SNAQ-SEQ
- Will addition of SNAQ-SEQ IS into plasma enable quantification of plasma variants?

- Spike 1000 or 10,000 IS in duplicate into 1 ml aliquots drawn from mixture of patient plasma retain
- MagMAX followed by Oncomine Pan-cancer liquid biopsy assay (+torrent server files modified for IS)
- ✓ Expected IS yields covaried with cfDNA (see poster TT33)
- ✓ Variant abundance levels ranged 100-3200 /ml plasma
- ✓ Good reproducibility (%CV < 36%)
- Next: measure ctDNA when varying cfDNA level in plasma

[1]Cancer Biol Ther. 2019; 20(8): 1057–1067

					ALT Counts				IS Coverage				Genome copies/ml plasma				
CHROM	POS	REF	ALT	VAF	10K	10K	1K	1K	10K	10K	1K	1K	10K	10K	1K	1K	%CV
chr17	37879588	A	G	30%	860	971	1059	1047	1992	2182	154	160	2159	2225	3438	3272	28
chr3	178952020	C	T	0.9%	28	30	24	36	1369	1397	120	101	102	107	100	178	17
chr7	55249063	G	A	32%	587	748	763	720	1494	1425	108	145	1965	2625	3532	2483	29
chr7	55259450	C	T	2.9%	95	92	133	126	1728	1951	112	162	275	236	594	389	36

SNAQ-SEQ enables measurement of ctDNA per ml plasma

* Study done in collaboration with Dartmouth Hitchcock Health System using Accukit™ ONCO1LB Catalog #1207

AccuGenomics

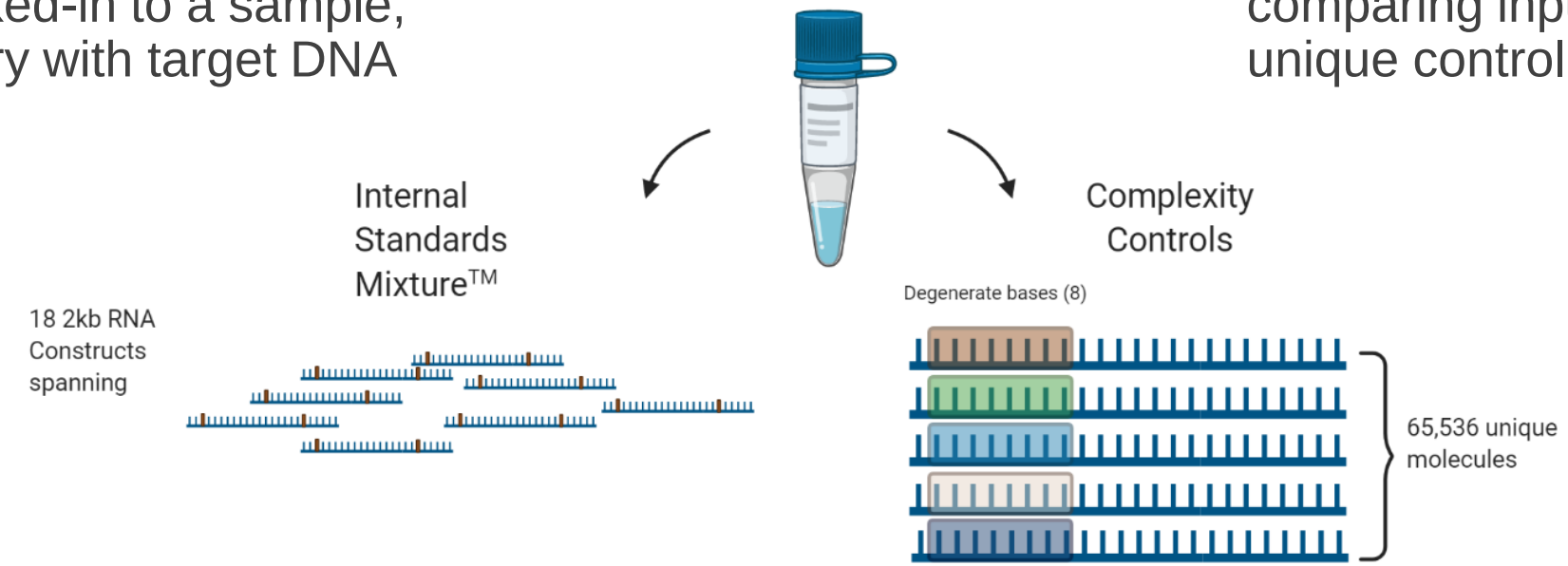
SNAQ-SEQ COVID-19 SCREENING

- ✓ Evaluate testing efficiency
- ✓ Measure viral load
- ✓ Adjust for reverse transcriptase artifacts as part of variant calling

AccuGenomics

WHAT IS A COMPLEXITY CONTROL?

Complexity Control (CC) is a degenerate sequence that, when spiked-in to a sample, will co-vary with target DNA



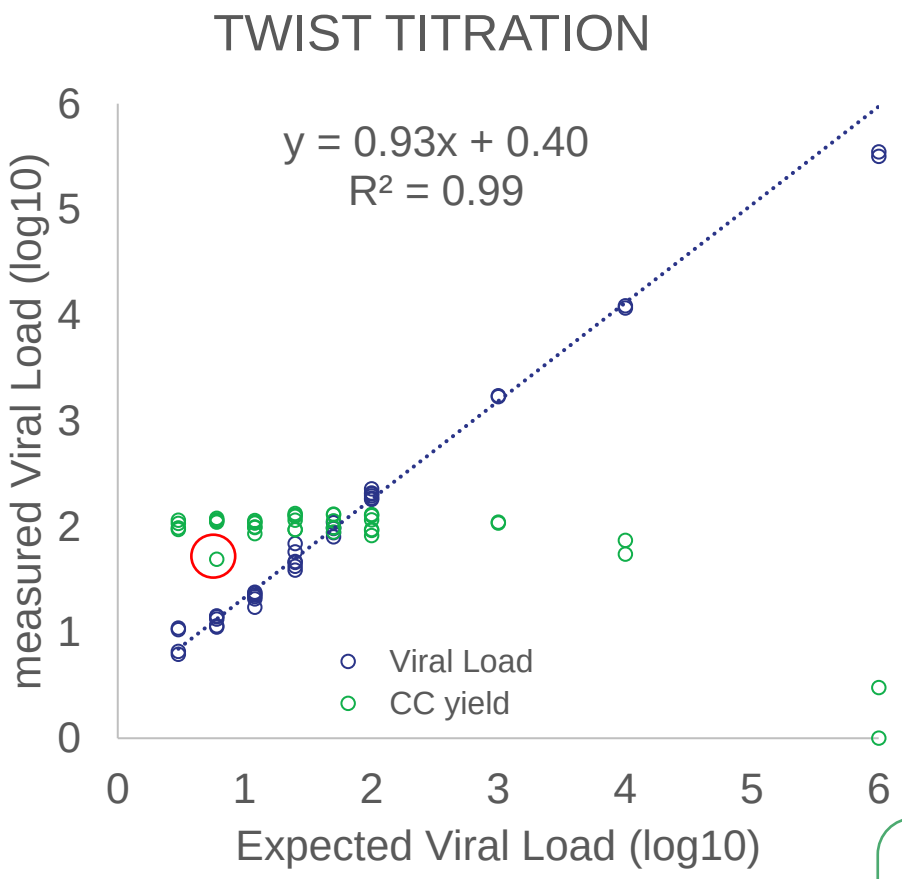
✓ CCs provide an estimate of library complexity capture by comparing input vs. detected unique control sequences

✓ Allows estimation of complexity loss due to deduplication

✓ Detects process drift that could impact results *before becoming significant*

AccuGenomics

COMPLEXITY CONTROL IN VIRAL TITRATION



- 1000 genome equivalents of Accukit™ SARS-CoV-2 IS added with master mix
- ✓ Indicated amount of TWIST COVID-19 RNA reference material (blue)
- ✓ Complexity capture was 20% (green)
- ✓ One sample showed 2-fold lower CC capture (red circle)
- Above 10^5 viral genome equivalents, the virus copies will outcompete 1000 controls for reads on flow cell

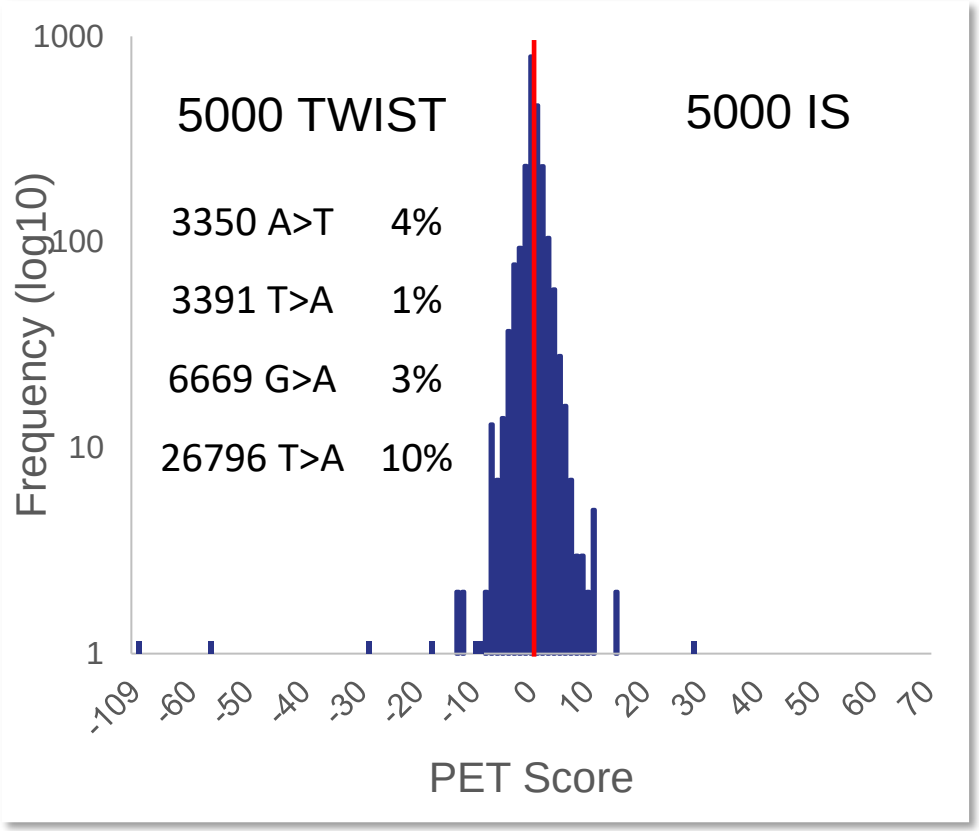
SNAQ-SEQ complexity control indicates template capture efficiency for every sample

Study done in collaboration with the Wellcome Sanger Institute using ARTIC v3 protocol with Accukit™ SARS-CoV-2 (v2, 1000) Catalog #1270

AccuGenomics

LIMIT OF BLANK ACCOUNTING FOR REVERSE TRANSCRIPTASE ERRORS

- Converting RNA to cDNA as part of ARTIC protocol creates random variants
- TWIST BioScience Synthetic RNA Reference Control compared spiked with AccuGenomics SARS-CoV-2 RNA Internal Standards
- ✓ SNAQ-SEQ RNA IS mirrors the sample RT error in the PET distribution
- ✓ The RT artifacts increase the limit of blank beyond 5% alpha
- User/software would make informed significance cutoff per sample

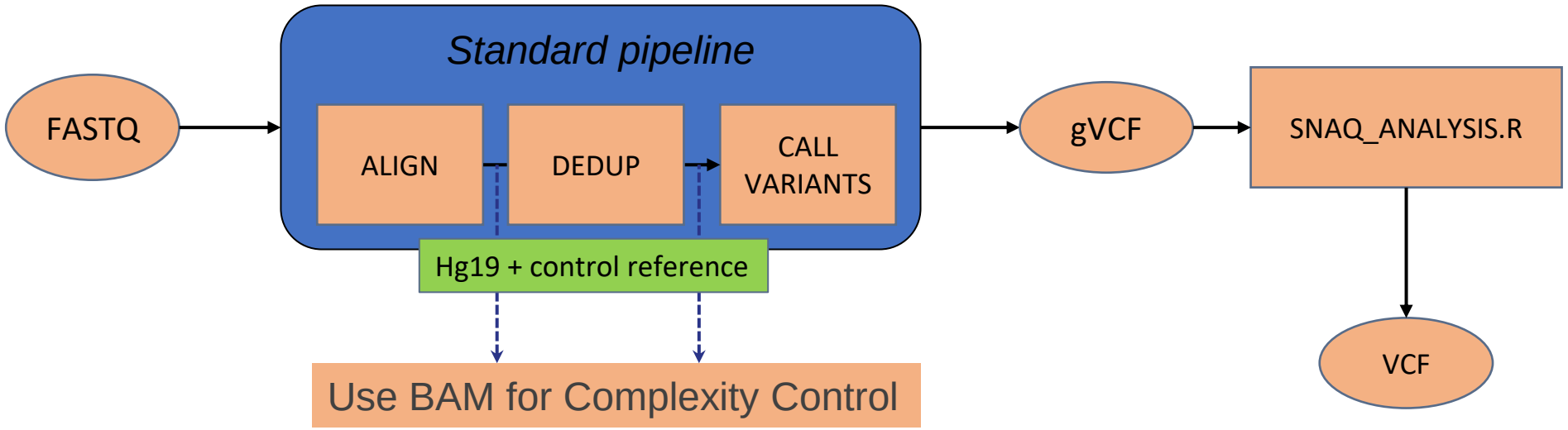


SNAQ-SEQ captures sequence noise generated from RT and enables low VAF calling in unoptimized pipelines

Study done in collaboration with the Wellcome Sanger Institute using ARTIC v3 protocol with Accukit™ SARS-CoV-2 (v2, 1000) Catalog #1270

AccuGenomics

SNAQ-SEQ Analysis Pipeline



SNAQ-SEQ is compatible with any pipeline that accepts reference genomes

AccuGenomics

SUMMARY

SNAQ-SEQ technology provides customized solutions for NGS assays across many platforms and delivers:

- ✓ Independent Quality Control for every variant in every sample
- ✓ QC for positives and negatives SNV
- ✓ Capture of RT sequence noise
- ✓ Better variant calling accuracy
- ✓ Potential to rescue poor quality samples

- ✓ Complexity control to measure template capture efficiency
- ✓ Concentration measurements of viral load and plasma ctDNA
- ✓ Compatibility with any pipeline that accepts reference genomes

ACKNOWLEDGMENTS

SEQC-2 Workgroup #2 Team

Joshua Xu
and MANY others

Baylor College of Medicine

Brian-Tyler St. Hilaire
Aviva Presser Aiden
Neva C. Durand
Namita Mitra
Saul Godinez Pulido

Birmingham University, UK

Joshua Quick

Dartmouth Hitchcock Health System

Sophie Deharvengt
Donald Green
Greg Tsongalis

NIIMBL Team (LOD Adventitious Agents)

Karen O'Connell (NCSU)
Caroline Smith-Moore (NCSU)
Peter Bernhardt (Celgene)
Dan Huang (Celgene
Bernice Westrek (Merck)
Veronica Fowler (Merck)
Melissa Scott (U Delaware)

The Wellcome Sanger Institute

Nicholas Redshaw
Julia Harvison
Naomi Park
Stephanie Lensing
Scott Goodwin

Sarah Cannon Molecular Diagnostics (CNV)

Kevin Balbi
Gareth Gerrard

AccuGenomics

ORDERING INFORMATION

Accukit Name	Catalog Number
SEQC2 Mix 4	1154
Accukit™ SARS-CoV-2 RNA (v2, 250)	1269
Accukit™ SARS-CoV-2 RNA (v2, 1000)	1270
Accukit™ BioContaminants	1306
Accukit™ ONCO1LB	1207
Accukit™ ONCO2ST	1208
Accukit™ Inherited Cancer CNV	1263

Contact Information:

Colette Saccomanno, PhD
Sales and Business Development
csaccomanno@accugenomics.com
201-893-2707
info@accugenomics.com

AccuGenomics
A higher standard of accuracy