Creating and Accurately Interpreting Clinical-grade Cancer Exomes with ACE Exome™

Contact: michael.clark@personalis.com

Michael James Clark, Elena Helman, Sean Boyle, Deanna Church, Mark Pratt, Gabor Bartha, Stephen Chervitz, Sarah Garcia, Shujun Luo, Jason Harris, Anil Patwardhan, Richard Chen, John West

Personalis, Inc. | 1350 Willow Road, Suite 202 | Menlo Park, CA 94025

Abstract

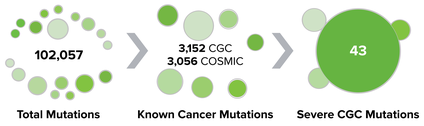
Exome sequencing is increasingly utilized to diagnose and direct treatment for cancer patients. However, there are many barriers to sequencing cancer exomes in an accurate and comprehensive way to assess all crucial cancer mutations. We have developed a comprehensive approach combining the Personalis ACE Exome™ and cancer analyses to improve accuracy and completeness of sequencing, call structural variants from exome, and comprehensively annotate cancer samples.

ACE Exome assesses major cancer genome mutations by augmenting coverage over greater than 1,200 cancer genes and over greater than 7,000 total genes of medical significance using additional optimized targeted enrichments on top of the standard exome. We also utilize altered sequencing protocols to fill in regions containing genomic elements that are typically hard to sequence by standard protocols. We compared ACE Exome to standard exome performance across an extended set of cancer genes. ACE Exome covered 99% of bases at greater than 25X depth with 16G of sequencing over 1,048 of 1,258 (83%) cancer genes. Standard exome platforms only covered 840 (67%), 763 (61%), and 676 (54%) cancer genes at this level.

We demonstrate ACE Exome utilized in three major clinical research modes. In the first, ACE Exomes were analyzed across a set of cancer cell lines, including 7 from the NCI-60 (A549 Lung, HT29 Colon, K562 Leukemia, MCF7 Breast, OVCAR-3 Ovarian, PC-3 Prostate, SK-MEL-28 Melanoma) and U87-MG Glioma, in order to demonstrate improved sensitivity for cancer mutations in well described cancer cell lines. In the second, a tumor/normal pair (prostate cancer) was analyzed via ACE Exome in order to demonstrate somatic analysis when a paired normal sample is available. In the third, a primary and metastatic tumor pair from the same renal cancer patient were assessed. These analyses demonstrate increased sensitivity for cancer mutations by ACE Exome analysis.

Improved Accuracy and Validation of Findings through Analysis of Cancer Cell Lines

A549 - Lung Adenocarcinoma

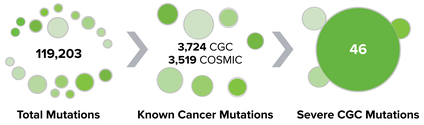


Gene	Effect	Condon Change	AA Change	Exon	Previously Described	COSMIC ID	NCI-60
MUTYH	non-synonymous	gGt / gAt	G368D	13	germline	n/a	yes
CBL	non-synonymous	Cgc / Tgc	R585C	11	somatic	1193372	yes
KRAS	non-synonymous	Ggt / Agt	G12S	2	somatic	517	yes
STK11	stop gained	Cag / Tag	Q37*	1	s/g	12925	no

A large number of total mutations are detected in a single unpaired analysis, but applying filters such as presence in known Cancer Gene Census (CGC) genes and COSMIC, low healthy population frequency, severe effect on coding, and actionability leads to a much more refined and meaningful list of mutations.

This list of mutations included variants previously detected by the NCI via exome sequencing, but also some not seen in their data. Of note in **A549**, we detected a homozygous somatic mutation in *STK11* (right) not detected in the NCI-60 exome, but which the NCI lists as a known cancer mutation in this cell line. It is not clear why the NCI exome missed this particular variant.

HT29 - Colorectal Adenocarcinoma



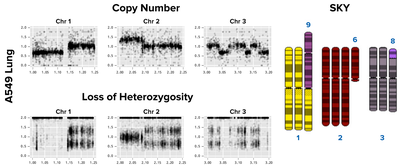
Gene	Effect	Condon Change	AA Change	Exon	Previously Described	COSMIC ID	NCI-60
TP53	non-synonymous	cGt / cAt	R141H	5	s/g	10660	yes
SMAD4	stop gained	Cag / Tag	Q311*	8	s/g	14163	yes
PIK3CA	non-synonymous	Cct / Act	P449T	8	somatic	18601	yes
ETV6	non-synonymous	Aga / Gga	R396G	7	somatic	1359370	yes
ABL1	non-synonymous	cCg / cTg	P810L	11	somatic		yes
APC	frameshift	gaa / gAaa	E1554E*	16	s/g	19718	no
APC	stop gained	Gag / Tag	E835*	14	s/g	18566	yes
BRAF	non-synonymous	gTg / gAg	V207E	6	somatic	476	yes
NF1	non-synonymous	gCa / gGa	A2218G	43	s/g		no

Gene	Drugs
ABL1	Dasatinib;Nilotinib;Bosutinib;Regorafenib;Ponatinib
BRAF	Sorafenib;Vemurafenib;Regorafenib;Dabrafenib

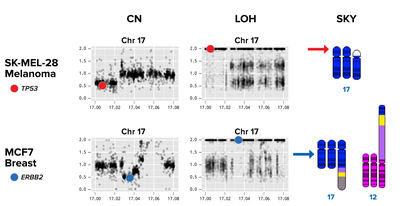
We consistently detected cancer mutations known to be present in these particular cell lines, and our annotation and filtering approach distilled out mutations in major driver genes as well as novel mutations such as the NS-SNP in *NF1* detected in the HT29 cell line.

We also used the Personalis annotation engine to annotate against gene-drug databases to identify cancer drugs associated with mutated genes we detect via ACE Exome (shown right). These annotations may help guide therapeutic decision-making when used in conjunction with clinical samples.

Structural Variant and Copy Number Results



Highly accurate large-scale copy number states were determined from ACE exome data by calculating normalized coverage-based copy number (CN) and loss of heterozygosity (LOH). These results are highly consistent with NCI-60 SKY data. In **A549**, a general triploid state is observed genome-wide, with evidence for many additional known del/dups.



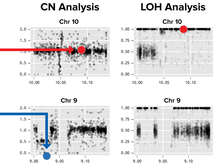
Other NCI-60 Cell Lines

Cell Line	Cancer Type	Total # Mutations	Severe CGC Mutations	Select Genes with Known Mutations	Select Genes with Unknown Mutations
K562	Leukemia	115,016	50	*FANCC, KLF4, PDGFRA, NOTCH1*	*TP53, MYB, RAD21*
MCF7	Breast	96,531	41	*ERBB2, BCR, PDGFB*	*PIK3CA, MET, MYH9*
OVCAR3	Ovarian	99,356	37	*TP53, JAK2, PIK3R1*	*PMS2, ERCC2, NOTCH1*
PC-3	Prostate	108,864	32	*TP53, FANCF, FAM46C*	*PCM1, KMT2C, FANCA*
SK-MEL-28	Melanoma	110,144	46	*PTEN, TP53, EGFR, BRAF, APC, NOTCH1*	*SEPT9, CREB3L2, PCM1, CHEK2,*

For each NCI-60 Cell Line analyzed, severe mutations were detected in a number of canonical cancer genes. Many of these were previously known to be present in these cell lines, but in each case, additional previously unknown severe mutations in cancer genes were detected.

U87-MG Glioma

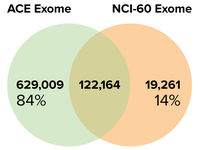
Gene	Mutation
PTEN	Homozygous splice site mutation
NF1	Compound heterozygous (ts/ns)
NOTCH2	Compound heterozygous (ns/ns)
KMT2C	Non-synonymous coding mutation
CDKN2A	Complete gene deletion

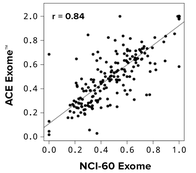


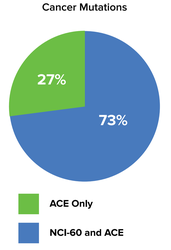
We compared ACE exome sequencing of U87-MG to published whole genome sequencing (WGS). Known severe mutations in cancer genes were confirmed. In addition, large-scale structural mutations detected by ACE exome were highly consistent with those detected by WGS.

ACE *vs.* NCI-60 Exome

We compared variant findings by ACE Exome with those found by the NCI exome tests on these seven cell lines. ACE Exome detected 86% of the variants detected by NCI's exome, but also a great number more (Venn Diagram). It is possible some of these additional variants are false positives, but we do see a highly consistent trend with regards to allele fraction between the two exome sets (scatter plot, r=0.84), suggesting consistency in variant calling accuracy. Greater than four times as many cancer mutations were called with ACE exome than NCI's exome (pie chart).



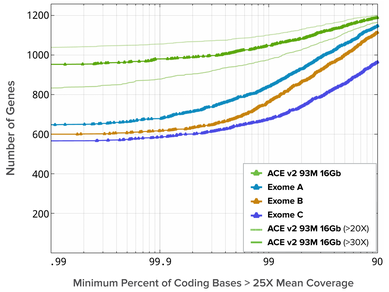




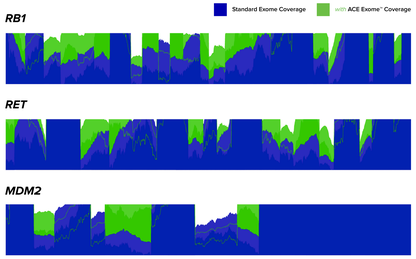
Personalis ACE Exome™ and Cancer Analysis

For each of the studies presented here, we performed **Accuracy and Content Enhanced (ACE) exome sequencing** to 12Gb of total sequence. ACE exome increases coverage over all biomedical regions of the genome including coding exons for over **7,800 genes**, all non-coding yet disease-associated variants (*e.g.*, intronic mutations and regulatory loci), and untranslated regions. This list of genes includes over 1,200 known cancer genes, including those on the Cancer Gene Census list and additional genes not yet on the Census list.

Finishing Analysis of 1,258 Cancer Genes by Assay



ACE Exome fills in regions with low average depth in a standard exome sequencing run as demonstrated in the plots below. These plots show (**in blue**) standard exome depth at 12G across the coding exons of three cancer genes. ACE Exome depth (**in green**), demonstrates that ACE tops up coverage over regions the standard exome poorly enriches. These increases in coverage lead directly to improved accuracy in variant calling in these regions.



We investigated completeness of coverage over a comprehensive list of 1,258 cancer genes (including the Cancer Gene Census list) by comparing the current ACE Exome with current standard exome offerings from three major commercial vendors, each run to a total of 16G of sequencing. **ACE Exome covered greater than 99% of bases at greater than 25X mean depth for 1,048 (83%) genes**. Standard exomes only covered 840 (67%), 763 (61%), and 676 (54%) respectively. Standard exomes were designed for technical specifications rather than clinical accuracy. ACE Exome sacrifices some efficiency and on-target enrichment for coverage of and accuracy across cancer genes.

Somatic Analysis: Primary Tumor and Paired Normal (Prostate Cancer)

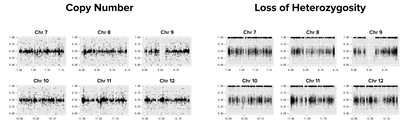
We performed ACE exome sequencing and analysis on a primary prostate tumor and paired normal sample from the same patient. The analytical steps taken differed from those performed on the cancer cell lines because paired somatic calling could be performed. Somatic variant calling with a paired normal sample is a very powerful approach that much more accurately reveals the somatic mutations in the tumor. It also allows for sensitive somatic mutation detection in subclones and lower cellularity samples.

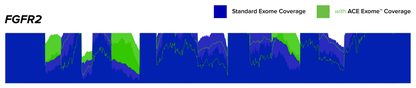
As might be expected, far fewer total somatic mutations were detected in the primary tumor paired with normal analysis than in the cancer cell lines analyses. After filtering these variants for presence in our cancer gene and variant databases, predicted effect, population frequency, and actionability, two somatic variants remained to be reported (TABLE 1). Neither of these variants were known cancer mutations, but mutations in *FGFR2* do have known therapies in multiple tumor types.

TABLE 1: Severe Somatic Cancer Mutations Detected in the Prostate Tumor

Variant	Gene	Effect	Codon Change	AA Change	Exon	Predicted Effect	Previously Described	Tissue Type: Frequency	Drugs
10:123258064 C>T	*FGFR2*	non-synonymous	aaG / aaT	K31N	4	deleterious	somatic	MEL:28|UCEC:28|LUAD:15|LUSC:11|CRC:8.6|MM:7.4|BRCA:6.8|ESO:3.7|BLCA:3.1|HNSC:3.1|KIRC:3.1|PRAD:3.1|DLBCL:2.5|GBM:1.9|OV:1.9|MED:0.62	Palifermin; Thalidomide; Regorafenib
16:23619219 T>C	*PALB2*	non-synonymous	Atg / Gtg	M1106V	12	neutral	germline	LUAD:22|LUSC:19|UCEC:18|MEL:16|CRC:14|HNSC:7.6|ESO:5.1|KIRC:5.1|OV:5.1|BLCA:3.8|BRCA:3.8|MM:2.5|AML:1.3|MED:1.3	

Much as we did not detect many severe somatic small variants in this tumor, it also appeared to have a completely normal copy number state with no large copy number variations on any chromosome.

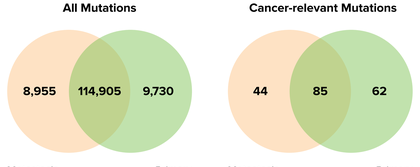




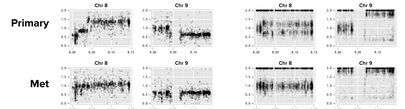
Certain *FGFR2* isoforms have variable expression changes in prostate tumors, and these differences in expression can correlate with prostate tumor progression. The somatic mutation identified has an unknown effect on expression. This stresses the importance of pairing expression analysis with exome analysis for a complete picture of tumor biology and progression. We have proceeded with expression studies on this particular tumor and are paying particular attention to *FGFR2* expression.

Tumor-only Comparison: Primary and Metastatic Tumors (Renal Cancer)

Using a similar analysis approach as the one we used for analyzing the cancer cell lines above, we performed ACE Exome sequencing and tumor-only analysis on a primary tumor and metastatic tumor from the same renal cancer patient. These Formalin-Fixed Paraffin-Embedded (FFPE) samples were processed using a custom DNA/RNA co-isolation from FFPE protocol so that we would be able to follow up our exome studies with transcriptome analysis from the same FFPE curls. After tumor-only analysis, we performed a combined analysis of the primary and metastatic tumor to understand how the met differed from the primary it was derived from and search for actionable targets present in the met that were not present in the primary.



There is a very high overlap between mutations detected in the primary and metastatic tumor samples, as shown in the Venn diagrams above. We did note slightly higher numbers of primary-specific mutations, likely owing to higher heterogeneity in the primary tumor. More surprising was the substantially higher number of primary- and met-specific cancer-relevant mutations. These could suggest that the metastatic tumor was derived from a subclone of the primary tumor lacking a number of cancer mutations, and that it picked up additional cancer mutations as it metastasized.



Structural differences were observed as well. For example, Chr 8 is triploid in the primary but diploid in the met. Meanwhile, 9p appears intact in the primary, but it is homozygous and duplicated in the metastatic tumor.

Gene	Effect	Predicted Effect	Drugs	Primary	Metastasis
ERBB2	non-synonymous	unknown	yes		0/1
NOTCH1	non-synonymous	deleterious		0/1	0/1
BAP1	non-synonymous	deleterious			0/1
HRAS	splice-site donor	unknown		0/1	0/1
HRAS	non-synonymous	deleterious		0/1	0/1
VTI1A	non-synonymous	unknown		0/1	0/1
BRCA1	non-synonymous	neutral		0/1	0/1
KMT2A	non-synonymous	deleterious		0/1	0/1
RET	non-synonymous	neutral	yes	0/1	
PIK3CA	non-synonymous	unknown			0/1
PIK3CA	non-synonymous	deleterious		0/1	
TSC2	non-synonymous	neutral		0/1	

Some mutations in major cancer genes were unique to either the primary or met. For example, a NS-SNP in *ERBB2*, a gene for which there are drug treatments, was unique to the met, suggesting a possible treatment path not present in the prmary. We also noted different mutations in *PIK3CA* in both primary and met. As with the structural variants (above), it is evident from the small variants that the met was derived from a subclone of the primary, but it displays numerous mutations unique to itself as well.

www.personalis.com | info@personalis.com | +1 855-436-6634 | +1 650-752-1300

POS-CAMC-1.0 May 2014

Personalis®
Pioneering Genome-Guided Medicine

Pioneering Genome-Guided Medicine