Integrating RNA/DNA analysis with a comprehensive cancer panel to improve interpretations of advanced kidney tumors

Sean Michael Boyle, Thai H. Ho, Michael James Clark, Alexander S. Parker, Ravi Alla, Elena Helman, Shujun Luo, Scott Kirk, Parvin Sripakdeevong, Mirian Karbleashvili, Deanna M. Church, Michael Snyder, Rich Chen, & John West

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

Contact: sean.boyle@personalis.com

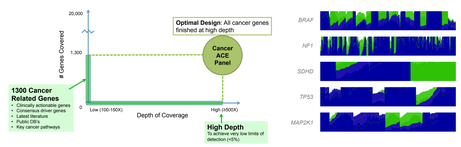
Sean Michael Boyle, Thai H. Ho, Michael James Clark, Alexander S. Parker, Ravi Alla, Elena Helman, Shujun Luo

Introduction

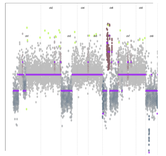
Next-generation sequencing is now commonly applied to measure genomic modifications, providing an accurate view of cancer progression. However, several challenges, including tumor purity, heterogeneity and the broad set of cancer associated cellular functions, increase the difficulty of interpreting cancers and make analyzing them distinct from analysis of germline diseases. The ability to detect low allelic mutations in the broad set of cancer-associated genes is critically important. To respond to those needs, we have developed the ACE Extended Cancer Panel, a targeted enrichment sequencing platform that includes over 1,300 cancer genes and 200 miRNAs. In this study, we applied our integrated RNA/DNA analysis and cancer panel platform to identify important driver mutations in pre and post treatment advanced renal cancers, a prevalent tumor type affecting more than 60,000 people in the United States each year.

Our goal with this study was to test the importance of integrating DNA and RNA analysis, which allows for detection of mutational events which cannot be observed by either method alone. With RNA analysis it is possible to detect gene fusion events, expression of small variants (SNVs/indels), and gene expression levels. DNA analysis allows for detection of loss of heterozygosity (LOH), copy number variants (CNVs), and small variants (SNVs/indels). By integrating both the RNA and DNA, it becomes possible to identify unique variant classes, such as unexpressed variants, allele specific expression, effects of copy number changes on gene expression levels, and how fusion gene integration sites affect copy number changes. More accurate interpretation of how variants affect important cancer pathways can also be assessed.

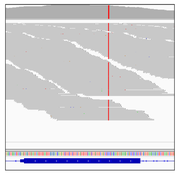
Analyzing Cancer Samples Using Ace Cancer Panel (DNA)



Applying Personalis Ace Cancer Panel To Finish Genes - The Ace Cancer Panel was specifically designed to increase both gene coverage and the ability to detect variants in important cancer genes.

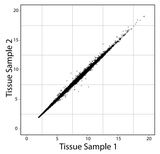


Copy Number Analysis - We apply cutting edge techniques to identify copy number change between tumor and normal tissue samples.

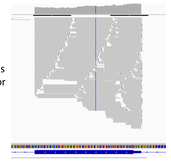


SNV / Indel Variant Calling - A carefully selected set of core tools call both germline and somatic variants in normal and/or tumor tissue samples. Identified variants are annotated from a very large set of genomic and pharmacogenomic databases.

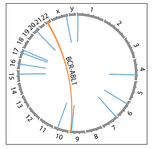
Analyzing Cancer Samples Using Ace Cancer Panel (RNA)



Gene Expression Analysis - We apply well known and thoroughly vetted approaches to calculate gene expression between either tumor/normal samples or groups of tumor samples.



SNV / Indel Variant Calling - A carefully selected set of core tools call both germline and somatic variants in normal and/or tumor tissue samples. Identified variants are annotated from a very large set of genomic pharmacogenomic databases.



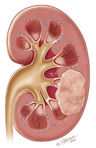
Gene Fusion Detection - Structural rearrangements can lead to functionally important gene fusions. We identify these fusions at base resolution and annotate them for improved downstream interpretation.

Advanced Renal Cell Carcinoma Samples Interpreted

Tumor Specifics:

Tumor Type: Renal Cell Carcinoma
Tumor Stage: 4
Sample Breakdown: 5 Patients Including 1 Xenograft & 1 Resistant

Molecular Classes Analyzed: RNA and DNA



Sequencing Specifics:

Sequencing Technique: Personalis ACE Cancer Panel For RNA and DNA
Sequencing Depth: 12G (DNA) & 3G (RNA)

Variant Identification Specifics:

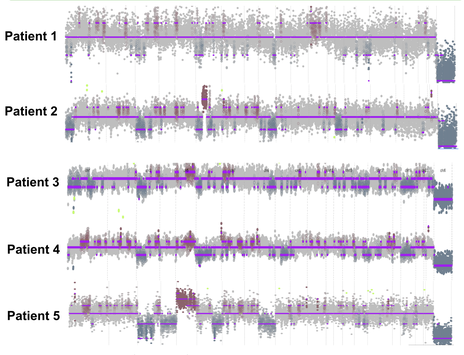
SNV Identificaiton DNA: Somatic Variant Detection (Paired Tumor-Normal)
SNV Identification RNA: Somatic Variant Detection (Paired Tumor RNA-Normal DNA)
Copy Number Alteration (CNA): Detection From Tumor DNA
Gene Fusion: Detection From Tumor RNA
Gene Expression: Gene Read Counts From Tumor RNA

A Comprehensive Analysis of VHL Mutations Across Patients

Individual	Copy Number Alteration	Small Variant	Small Variant Type	Allele Specific	DNA %	RNA %
Patient 1	Yes	Yes	Frame Shift	Yes	5	0
Patient 2	Yes	No				
Patient 3	Yes	No				
Patient 4	Yes	Yes	Non Synon	No	31%	25%
Patient 5	Yes	Yes	Frame Shift	No	50%	37%

VHL mutations are major contributors of RCC progression and are biallelic in several patients - VHL is a major driver of sporadic RCC, with >90% of cases containing mutations in this gene. VHL mutation usually follows a two hit profile, wherein initially a single mutation forms, such as a chromosomal arm loss, followed by a second independant mutation, such as a small variant, leading to complete innactivation. It has been demonstrated that VHL with as little as one additionally mutated gene is sufficient for RCC progression. Along these lines, all of our patients contain a single copy chr3 deltion and three of our patients contain additoanl small variants in VHL.

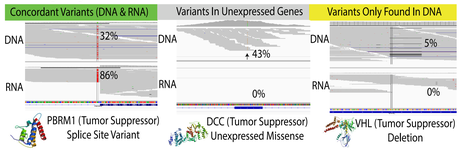
Copy Number Alterations (CNAs) Display 3p Deletion



The majority of patients display a single copy chromosome 3p deletion - This deletion, in addition to the 5p amplification observed in patients 3, 4, and 5, are characteristic of RCC. The 3p deletion houses the well known RCC gene VHL, which is commonly mutated or deleted. The chr 5q amplification houses EZH2 (histone modification), STC2 (stress response), VCAN (cell adhesion and migration), and mTOR regulator all of which are well known to be highly involvedin RCC progression.

Detecting RCC Small Variants (SNVs/Indels) In RNA and DNA

Integrating Small Variants Identified in RNA and DNA Results In Three Distinct Classes



Integrated Analysis Allows for Increased of Variant Confidence - By identifying SNVs and indels in both RNA and DNA we were able to better asses small variants: increasing confidence in concordant variants (Observed in both DNA & RNA) and decreasing confidence in unexpressed variants (DNA only calls or variants in unexpressed genes).

High Impact Variants Identified in Each of The Five Samples

Patient 1

Gene	Variant Class	Variant Position	Ref	Alt	DNA%	DNA Coverage	RNA %	RNA Coverage	Integrated Class
VHL	Frame Shift	Chr3:10183681	AGGAGGAGA TGGAGGCCG GGCGGCCG	A	5%	1,050	0%	80	Allele Specific
PIK3CB	Stop Gained	Chr3:138417823	G	A	2%	460	0%	233	Allele Specific
RET	Non Synon Coding	Chr10:43601980	G	T	10%	630	NA	0%	Coverage

Patient 2

Gene	Variant Class	Variant Position	Ref	Alt	DNA%	DNA Coverage	RNA %	RNA Coverage	Integrated Class
NF2	Frame Shift	Chr22:30035135	AT	A	29%	330	24%	303	Concordant
ERBB2	Non Synon Coding	Chr17:37881152	G	T	15%	692	31%	368	Concordant
KMT2A	Non Synon Coding	Chr11:118344558	A	G	2%	443	0%	514	Allele Specific
XPO1	Stop Gained	Chr2:61719752	C	A	19%	488	16%	3,367	Concordant

Patient 3

Gene	Variant Class	Variant Position	Ref	Alt	DNA%	DNA Coverage	RNA %	RNA Coverage	Integrated Class
TET2	Stop Gained	Chr4:106158275	C	G	21%	596	20%	610	Concordant
PBRM1	Splice Site Donor	Chr3:52668616	A	T	32%	334	86%	125	Concordant - ASE
SETD2	Frame Shift	Chr3:47058700	AT	A	27%	527	27%	746	Concordant
GNAS	Non Synon Coding	Chr20:57429312	C	T	26%	2,778	0%	0	Allele Specific
AR	Codon Deletion	ChrX:66766356	TGGC	T	6%	743	NA	NA	Coverage

Patient 4

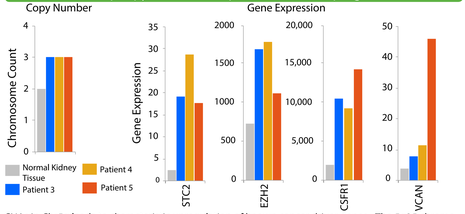
Gene	Variant Class	Variant Position	Ref	Alt	DNA%	DNA Coverage	RNA %	RNA Coverage	Integrated Class
VHL	Non Synon Coding	Chr3:10188215	A	G	31%	534	25%	2,075	Concordant
TP53	Non Synon Coding	Chr17:7578181	G	A	10%	357	12%	670	Concordant
NFE2L2	Non Synon Coding	Chr2:178098799	T	G	5%	494	17%	7,454	Concordant
NOTCH1	Non Synon Coding	Chr9:139400110	C	T	1%	735	1%	233	Concordant
BCR	Codon Deletion	Chr22:23653925	TGGA	T	9%	2,104	0%	754	Allele Specific
SETD2	Frame Shift	Chr3:47147603	A	C	8%	202	11%	923	Concordant

Patient 5

Gene	Variant Class	Variant Position	Ref	Alt	DNA%	DNA Coverage	RNA %	RNA Coverage	Integrated Class
VHL	Frame Shift	Chr3:10191544	CA	C	50%	200	37%	307	Concordant
TET1	Frame Shift	Chr4:106193856	C	CG	45%	305	16%	161	Concordant - ASE
PBRM1	Frame Shift	Chr3:52649468	TA	T	47%	76	4%	320	Concordant - ASE
NPM1	Non Synon Coding	Chr5:170818423	C	A	20%	715	30%	22,048	Concordant

Important RCC driver variants were found in all patients - High impact somatic variants were identified in each patient and VHL mutations, which are important divers in RCC, were observed in patients 1, 4, and 5. While these VHL non synonymous coding and frame shift variants in patients 4 and 5 were observed in similar ratios in both the DNA and RNA, the low allelic ratio (5%) frameshift deletion in patient 1 was only observed in the DNA, suggesting that this variant may in exhibit allele specific expression. Additionally, we observed PBRM1 variants in patients 3 and 5 and NF2 in patient 2. As a whole we observe good concordance between small variants observed in RNA and DNA.

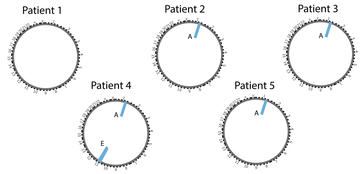
Patients with 5q Copy Number Amplification And Upregulation



CNAs in Chr5q leads to characteristic upregulation of known cancer driver genes - The 5q35 chromosomal region is often amplified in RCC. This region, which contains 60 genes, houses a number of genes which are known to be upregulated in RCC as a result of this cCNA. In each of the patient samples with amplifications in 5q, we observe amplifications in well known RCC driver genes including EZH2 (histone modification), STC2 (stress response), VCAN (cell adhesion and migration).

RCC Progression Likely Not Resulting From Fusion Genes

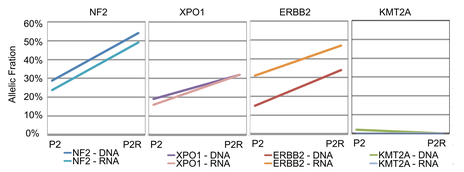
Fusion Event	Gene 1 symbol	Gene 2 symbol	Fusion description	Gene 1 Fusion Site	Gene 2 Fusion Site
A	METTL13	DNM3	< 100 kbp distance	1:171783163:+	1:171890888:+
B	RNASET2	RPS6KA2	readthrough, < 100 kbp distance	6:167352383:-	6:167271747:-
C	IMMP2L	DOCK4	readthrough,< 1,000 kbp distance	7:111127294:-	7:111462512:-
D	ACCS	EXT2	readthrough, < 100 kbps distance, healthy	11:44102670:+	11:44129233:+
E	AHNAK	MALAT1	no_protein_produc ,lincrna	11:62223188:-	11:65268472:+



Fusion genes are not major drivers of RCC - We did not observe any strong candidate fusion gene drivers in our RCC samples. As many RCC samples have been sequenced and analyzed yielding representatively few putative cancer driver gene fusions, this is not unexpected. Each four potential fusion events has been described as being directly downstream of each other or potentially being involved in readthrough events by our fusion gene identification pipeline.

Patient With NF2 Frame Shift Tumor Does Not Respond to Treatment

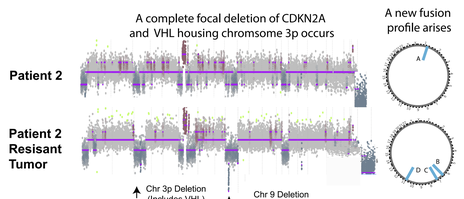
Gene	Patient 2	Patient 2 Resistant	Variant Class	Variant Position	Ref	Alt	P2 DNA%	P2 Resist DNA %	P2 RNA %	P2 Resist RNA %	Integrated Class
NF2	Y	Y	Frame Shift	Chr22:30035135	AT	A	29%	54%	24%	49%	Concordant
ERBB2	Y	Y	Non Synon Coding	Chr17:37881152	G	T	15%	34%	31%	47%	Concordant
KMT2A	Y	N	Non Synon Coding	Chr11:118344558	A	G	2%	0%	0%	0%	Allele Specific
XPO1	Y	Y	Stop Gained	Chr2:61719752	C	A	19%	32%	16%	32%	Concordant



Variants from primary which dissapear in the post treatment resistant tumor:
SLIT2 (2%), PIL1L1 (9%), KMT2A (2%), CHD4 (4%), NIN (11%), PSMO11 (2%)

Patient does not respond to mTOR inhibitor - While the other patient samples were clear cell, patient 2 was a mucinous tubular and spindle cell carcinoma. This patient did not contain any VHL or PBRM1 mutations. Instead, we observed a NF2 frame shift mutation. It has been demonstrated that RCC tumors of this type can be treated with mTOR inhibitors (rapalogs). Along these lines, our patient was treated with pazopanib and temsirolimus. However, the tumor conitinued to agressively grow.

In analyzing the posttreatment resistant tumor, we observe a dramatic increase the allelic ratios of cancer driver genes identified in the pretreatment sample. Additionally, we do not observe new high impact variants arising post treatment.

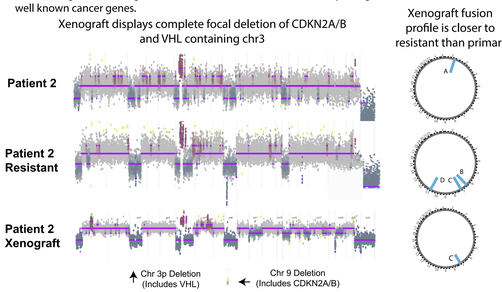


Structural rearangemnt suggests drug induced bottleneck effect - In the post treatment tumor sample the majority of tumor cells contain full single copy chromosomal deletions characteristic of RCC. For example, chr3, which houses the VHL gene. Additionally, we observe a focal two copy deletion involving the well known chr9p21 region, which houses CDKN2A, ARF, and CDKN2B.

A Mouse Xenograft Model Mimics Resistant Tumor

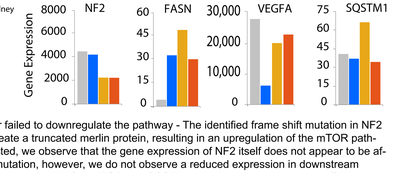
Gene	Patient 2	Patient 2 Recurrene	Patient 2 Xenograft	Variant Class	P2 DNA%	P2 Recuc DNA %	P2 Xeno DNA %	P2 RNA %	P2 Recur RNA %	P2 Xeno RNA %	Integrated Class
NF2	Y	Y	Y	Frame Shift	29%	54%	98%	24%	49%	97%	Concordant
ERBB2	Y	Y	Y	Non Synon Coding	15%	34%	55%	31%	47%	52%	Concordant
KMT2A	Y	N	Y	Non Synon Coding	2%	0%	1%	0%	0%	0%	Allele Specific
XPO1	Y	Y	Y	Stop Gained	19%	32%	46%	16%	32%	30%	Concordant
KRAS	N	N	Y	Non Synon Coding	0%	0%	12%	0%	0%	18%	Concordant
RET	N	N	Y	Non Synon Coding	0%	0%	2%	0%	0%	0%	Gene Unexpressed
RB1	N	N	Y	Non Synon Coding	0%	0%	51%	0%	0%	56%	Concordant
NKX2-1	N	N	Y	Frame Shift	0%	0%	44%	0%	0%	0%	Gene Unexpressed

Cancer driving small variants are most prominent in the xenograft model - We created a xenograft model from the primary tumor of patient 2. Interstingly, we observe that the allelic ratios in the mouse xenograft model follow the same trend we observed in the posttreatment resistant tumor patient. In fact, allelic ratios for these well known cancer driver genes increase greatly over both the primary and resistant tumor, indicating that this xenograft model is selecting for small variants in a similar fashion to the patient treatment. Additionally, we observe select new mutations arising only in the xenograft, including RET and KRAS, which while not commonly being mutated in RCC, are well known cancer genes.



Structural variants in the xenograft model most closely resemble the resistant tumor - As with small variants, we observe that structural variants in the xenograft model more closely resemble the posttreatment sample that the pretreatment primary.

Potential Reason For Treatment Failure In Patient 2



mTOR inhibitor failed to downregulate the pathway - The identified frame shift mutation in NF2 is believe to create a truncated merlin protein, resulting in an upregulation of the mTOR pathway. As expected, we observe that the gene expression of NF2 itself does not appear to be affected by the mutation, however, we do not observe a reduced expression in downstream mTORC1 complex targets FASN, VEGFA, or SQSTM1. This combined with the ineffectiveness of mTOR inhibitor treatment in the patient sample, suggests that the drug may not have have inhibited in the patient, leading to tumor progression and treatment resistance. Without RNA, we would not have been able to identify this potential resistance cause.

Summary

By applying both our RNA/DNA integration approach and our ACE Extended Cancer Panel to investigate advanced kidney cancer samples, we were able to classify many genomic alterations that drive cancer progression. For example, we found that anywhere from 5% to 30% of small variant mutations in driver genes called in the DNA reside in genes that are entirely unexpressed in the RNA. Likewise, we noted multiple cases of allele-specific expression, where high profile variants observed in DNA were entirely absent in the RNA despite high expression levels. Strong correlations between CNVs observed in DNA and gene expression changes found in RNA were also detected. We also observed that about half of the kidney tumors we tested were biallelic for VHL mutations, facilitating kidney cancer progression and others had PBRM1 or NF2 mutations. We also found that the resistant tumor from patient 2 continued increasing in allelic ratios in important cancer genes and, most importantly, that this treatment failure may have resulted from lack of inhibition in the mTOR pathway.

This analysis demonstrates how combining an integrated RNA/DNA approach with a cancer focused augmented enrichment panel allows for detection of both low allelic representation variants and unique variant classes, both of which are critical for accurate interpretation of cancer samples.

Personalis®
Pioneering Genome-Guided Medicine