Prioritizing germline and somatic variants potentially associated with papillary renal cell carcinoma in the p1RCC

Authors

In-Hee Lee, PhD
in-hee.lee@childrens.harvard.edu
Computational Health Informatics Program
Boston Children's Hospital, Boston, MA 02115

Sek Won Kong, MD
sekwon.kong@childrens.harvard.edu
Computational Health Informatics Program
Boston Children's Hospital, Boston, MA 02115
Department of Pediatrics
Harvard Medical School, Boston, MA 02115

Abstract

Prioritizing germline and somatic mutations associated with cause and progress of cancers is challenging due to false discoveries. We use open source variant prioritization tools to interactively filter variants. Then we cross-checked TCGA, ICGC, COSMIC and CENSUS databases to further prioritize papillary renal cell carcinoma associated genes.

Software used

gNOME (http://gnome.tchlab.org) and
gNOME-Report (http://gnome.tchlab.org/report/report_work.cgi/)

Reference: Lee IH, Lee K, Hsing M, Choe Y, Park JH, Kim SH, Bohn JM, Neu MB, Hwang KB, Green RC, Kohane IS, and Kong SW. Prioritizing disease-linked variants, genes, and pathways with an interactive whole-genome analysis pipeline. Hum Mutat. 2014 May;35(5):537-47. PMID:24478219.

Process

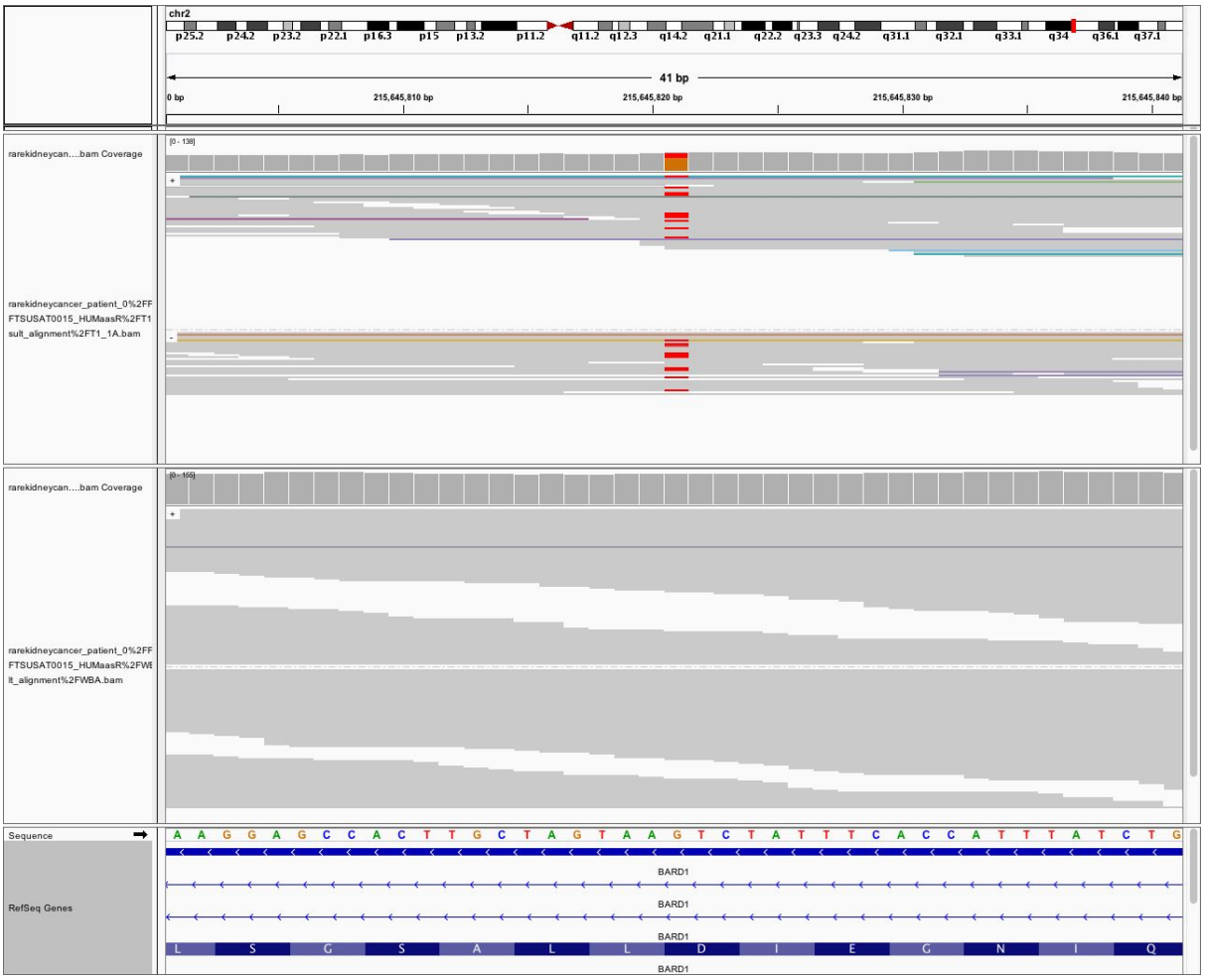
We used the provided VCF files prepared from germline and somatic WGSs. Two files are uploaded to the gNOME pipeline to annotate and filter variants according to following steps.

1. Select rare (AF < 1%) non-synonymous variants only found in tumor samples: found in T1_1A.snps.vcf/T1_1A.indels.vcf && not found in WBA.snps.vcf/WBA.indels.vcf [448 variants]
2. Intersect with the provided somatic mutation calls (supposedly used Strelka2) (gs://p1rcc/somatic/results/variants/somatic.snvs.vcf.gz; gs://p1rcc/somatic/results/variants/somatic.indels.vcf.gz) [46 variants]
3. Cross-compared with all somatic mutation from kidney cancers (source: cBioPortal) & Cancer Gene Census.
4. Annotate with PCGR (Personal Cancer Genome Reporter [https://github.com/sigven/pcgr])
    a. 8 tier-3 mutations (uncertain significance)
    b. 38 tier-4 mutations (other coding mutations)
    c. No higher tier mutations was reported.

Results

Top hit somatic mutation: *BARD1* missense mutation (NC_000002.11:g.215645821G>T, p.Asp259Glu).

- The mutant allele fraction (fraction of T allele) is 19/61 = 31.15%.
- *BARD1* is tumor suppressor gene among Cancer Gene Census (tier-1; i.e., genes with documented activity relevant to cancer).
- Mutations on *BARD1* gene are found among 1% of kidney cancer samples.
    - No mutation is found among TCGA renal papillary cell carcinoma
- *BARD1* gene on RefSeq: encodes a protein with interacts with the N-terminal region of *BRCA1*. The *BARD1/BRCA1* interaction is disrupted by tumorigenic amino acid substitutions in *BRCA1*, implying that the formation of a stable complex between these proteins may be an essential aspect of *BRCA1* tumor suppression.
    - Remark: the patient carries a common *BRCA1* germline variant (rs1799966, minor allele frequency: 34.98%)
- We confirmed the somatic mutation in *BARD1* in IGV as shown below.

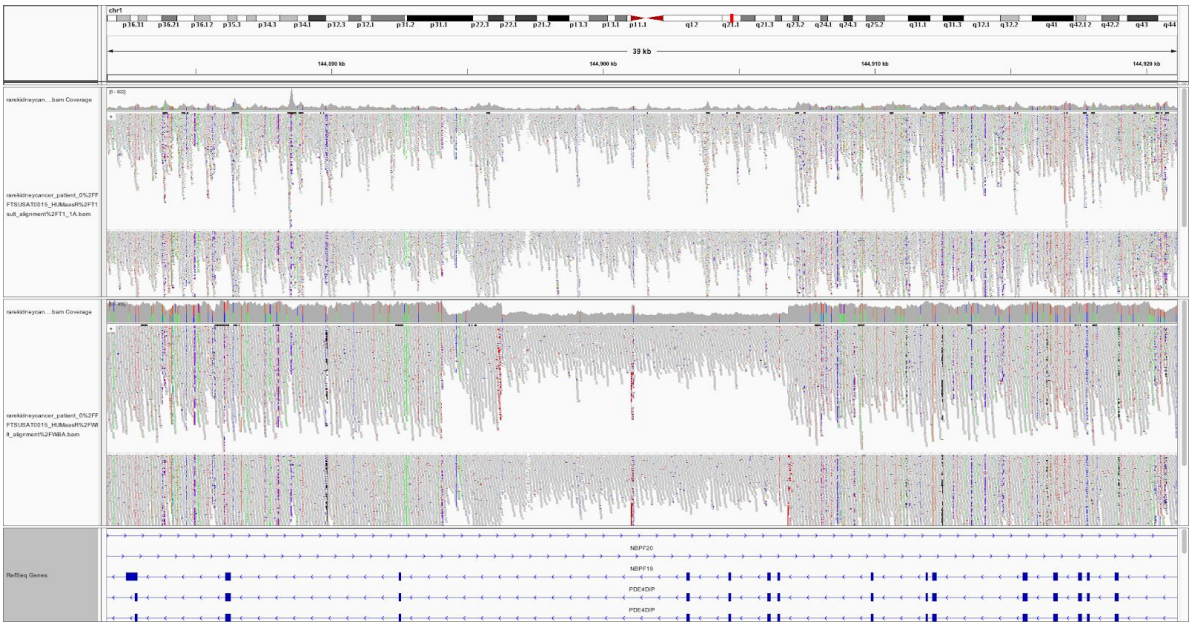


Other genes with tier 3 somatic mutations include:

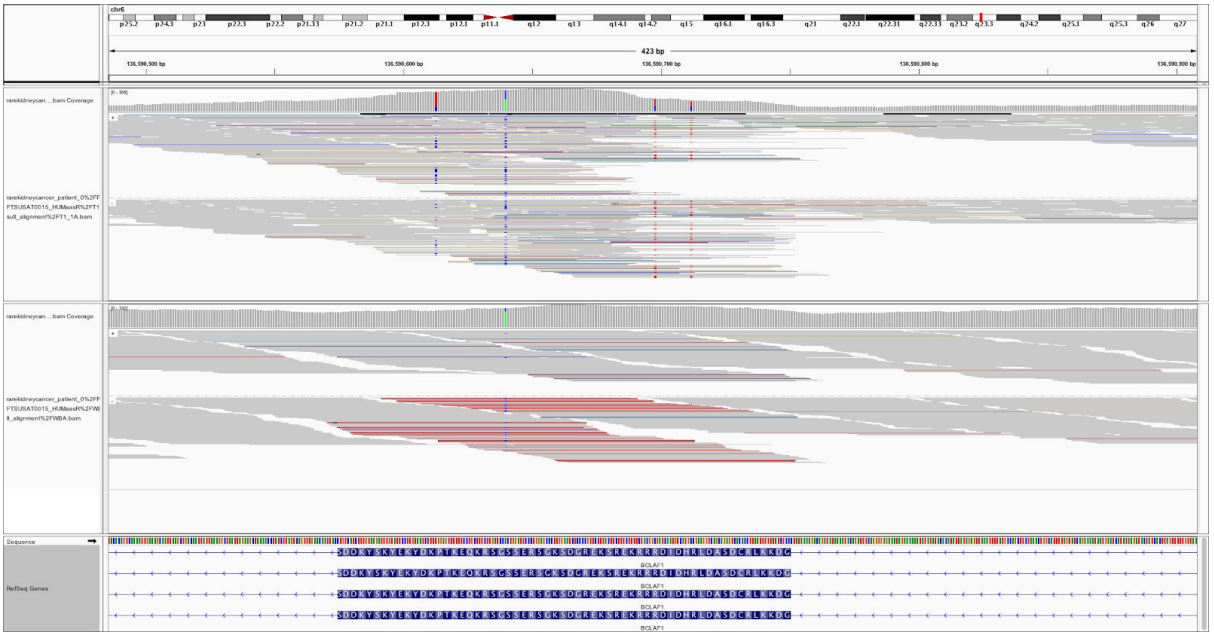
- *AHNAK, RBMX, ANAPC1, STARD13, DNAJ27, SCYL1*

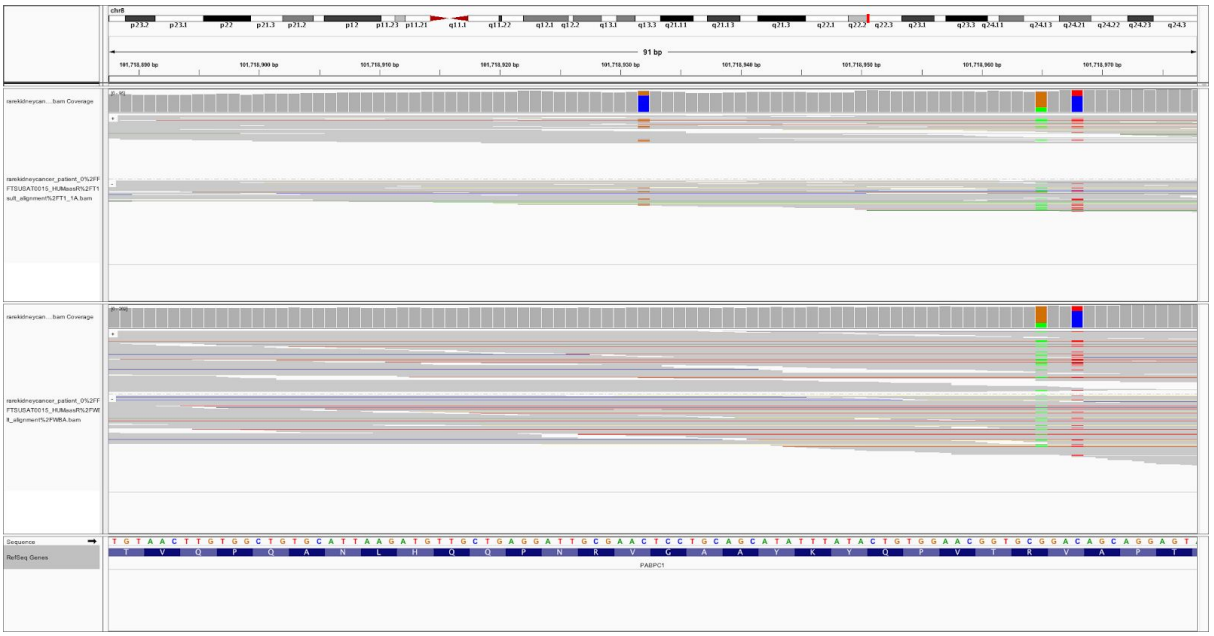
Interesting germline CNV observation in *PDE4DIP* gene

- *PDE4DIP* gene accumulated many germline coding variants (more than 10 variants). We suspected structural variant spanning *PDE4DIP* locus. Visually inspected BAM files and found ~20kb copy number loss in germline sequencing.



All the other candidates seemed to be false discoveries. Belows are two likely true positive examples in *BCLAF1* and *PABPC1*.




chr8
91 bp
rarekidneycan....bam Coverage
rarekidneycancer_patient_0%2FFFTSUSAT0015_HUMaselR%2FT1st8_alignment%2FT1_1A.bam
rarekidneycan....bam Coverage
rarekidneycancer_patient_0%2FFFTSUSAT0015_HUMaselR%2FWBst8_alignment%2FWBA.bam
Sequence
RefSeq Genes
PABPC1