Extended Abstract

A typical tumor has thousands of genomic variants, yet very few of these variants are thought to drive tumor growth. The remaining variants, termed passengers, represent the majority of all variants in cancer, but their functional role is poorly understood. *Functional* passenger variants are those that impact gene expression or activity, and while some of these changes are irrelevant, others may promote or inhibit tumor cell growth and survival, as is the case in *latent driver variants* and *deleterious passenger variants*, respectively. In this work, we explore the landscape of functional passenger variants in various cancer by leveraging extensive pan-cancer variation data from ~2500 uniformly processed whole cancer genomes. More specifically, we annotate and evaluate the impact of each variant, including SNVs, INDELs and SVs in the PCAWG dataset. Subsequently, we integrate the annotations and impact scores to quantify the overall burdening of various functional elements in cancer genomes. Furthermore, we show how overall functional burdening correlates with age at cancer diagnosis, patient survival time, and tumor clonality. This suggests a role of for *functional passenger variants* in cancer progression.

Inspection of variant functional impact scores across cancer cohorts, reveals that variants can be broadly classified into three distinct subgroups. The upper and the lower extreme of scored variants fall under traditional definitions of high impact driver variants and neutral passenger variants, respectively. In contrast, the intermediate functional impact regime comprises of *functional passenger variants*, which can further influence cancer progression by acting as latent drivers or through large burdening of various functional elements of the genome. Interestingly, we don't observe such splitting in germline variant impact score distribution. Consequently, we comprehensively analyzed the overall burdening of various genomic elements including TF (transcription factor) motifs in the PCAWG dataset. Presence of a variant in TF binding site can lead either to creation or destruction of TF motifs. In both cases, we observe differential burdening of *functional variants* among different cancer cohorts. Furthermore, the selective enrichment/depletion of *functional variants* in different TF motifs highlight distinct alteration profile associated with different component of regulatory networks.

Moreover, we integrated sub-clonal information along with functional burdening measure to investigate the role of high impact passenger variants in cancer evolution. We observed that high functional impact passenger variants along with LOF variants have a higher allelic frequency and a higher prevalence in parental subclones, signifying their important role in the early phases of cancer progression. We next sought to examine whether these functional mutations might exert a clinically meaningful impact on tumor progression. Therefore, we performed survival analysis to see if functional mutation burden predicted patient survival within individual cancer subtypes. We found that functional mutation burden predicted substantially earlier death in CLL and substantially prolonged survival in RCC, respectively. These results lend support to the hypothesis that functional passenger mutations are clinically meaningful, at least in some cancer subtypes. More specifically, the results suggest that latent drivers are more important than deleterious passengers in CLL but that the situation is reversed in RCC. In conclusion, our work highlights an important subset of the coding and noncoding variants that

originally identified as “passenger variants”, nonetheless show biologically and clinically relevant functional roles across a range of cancers.

*** passenger burdening is obviously not uniform: diff parts of the regulatory network are burdened differently by passengers in cancers**

*** the distribution of functional impact is not simple... there appears to be subgroups, one of which could correspond to latent drivers .... this is very different the functional impact in for germline**

*** we can divide passengers into high and low impact... if we do this, in general, we can find that there are more high impact passengers in the early passengers than the late**

*** furthermore, if we divide a given cancer, into two parts, one with lots of high impact passengers and another without, we can find differential survival**