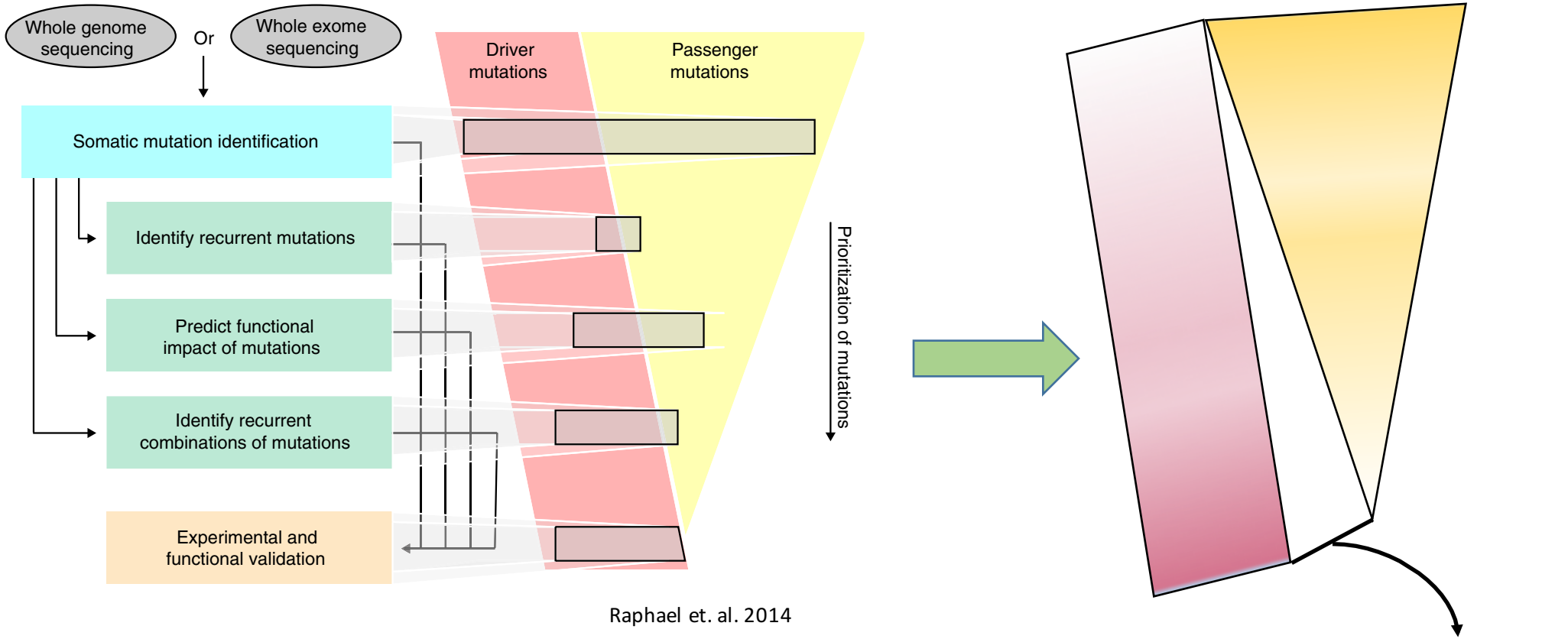
P2 Variation

05/20/2016

Research paper on functional impact of variants

- Develop comprehensive and consistent functional annotations for germ-line and somatic variants
- Using these determine functional impact of somatic v germline mutations in an individual, providing an overall statistical view of the functional "burdening" of a diseased individual.
- (Focus here is on a single individual and functional burden of mutations, including mutations with deleterious impact)
- Relation to TF binding sites & enhancers (gain or loss of function). Include TF families besides individual elements.
- (Might also include coding genes and ncRNAs for comparison)
- Relation to regulatory (ie TF binding) hubs and bottlenecks
- Relation to other genomics signals (eg replication timing and histone marks) and mutational signatures
- Relation to downstream gene expression for a mutation

Identifying driver mutations in cancer genome

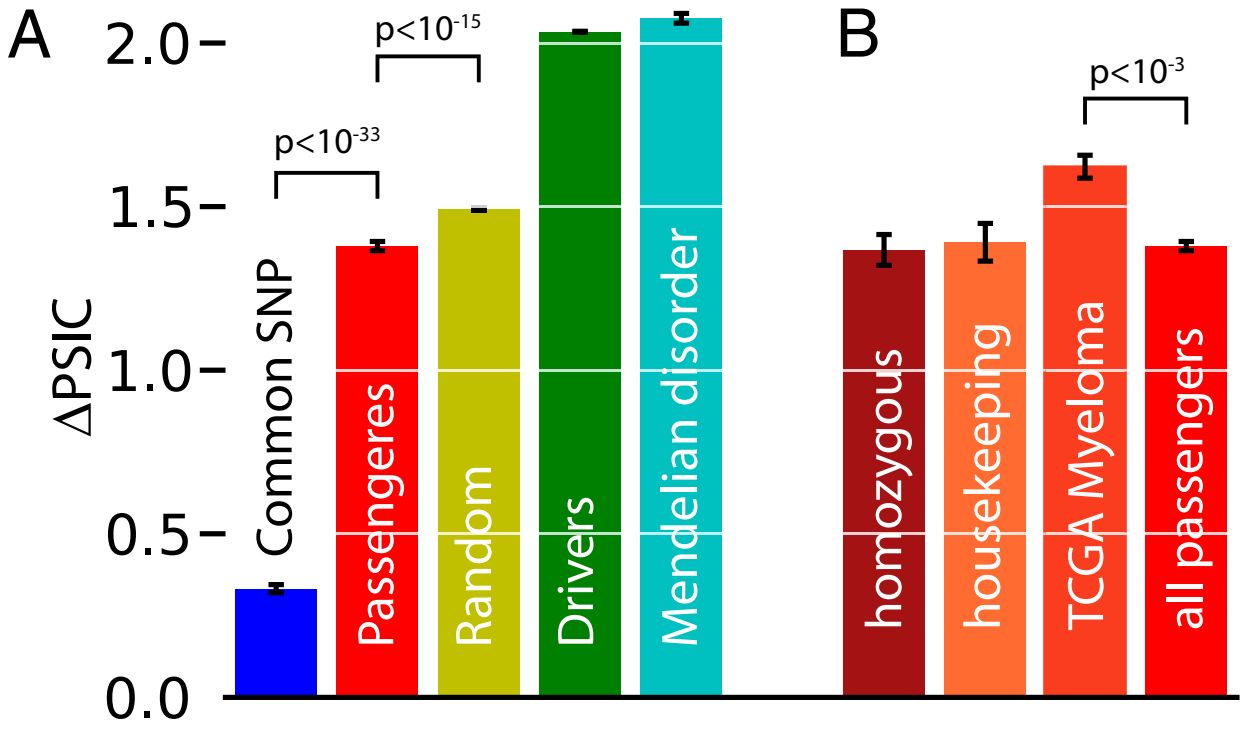


Raphael et. al. 2014

Majority of methods discretize somatic mutations as Driver and Passenger.

Intermediate regime between Driver & Passenger mutations
"mini-Driver" , "latent-Driver" , "deleterious Passenger"

Deleterious passenger mutations in cancer genome

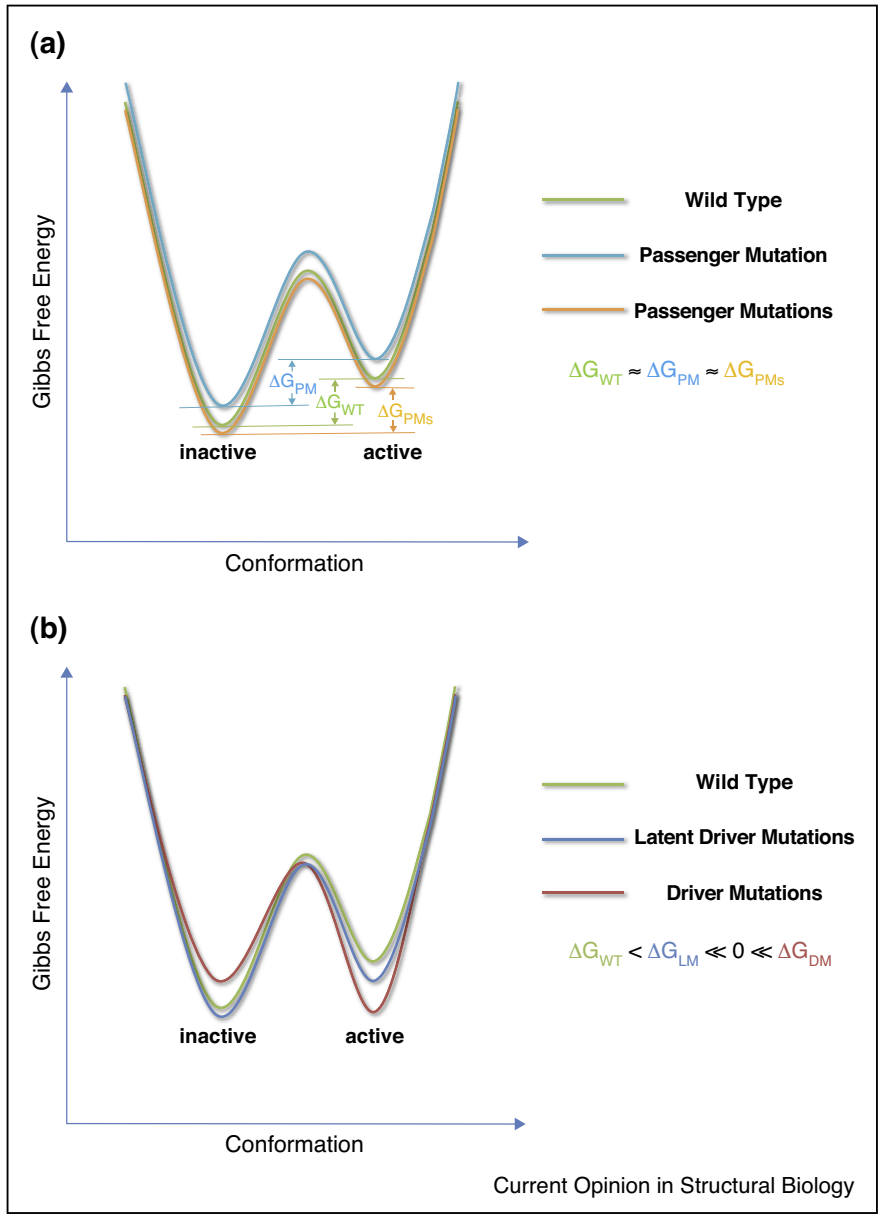


McFarland et. al. 2013

Polyphen score of passenger mutations is higher compared to common SNPs but closer to random mutations indicating presence of deleterious passenger mutations, which evades purifying selection.

Passenger mutations in house keeping genes also have high impact score suggesting presence of damaging passenger mutations is not only restricted to unimportant genes.

Latent Driver mutations in the coding region of the genome



Hypothesis : Latent Driver mutations are "passenger mutations", which can additively alter the protein folding profile.

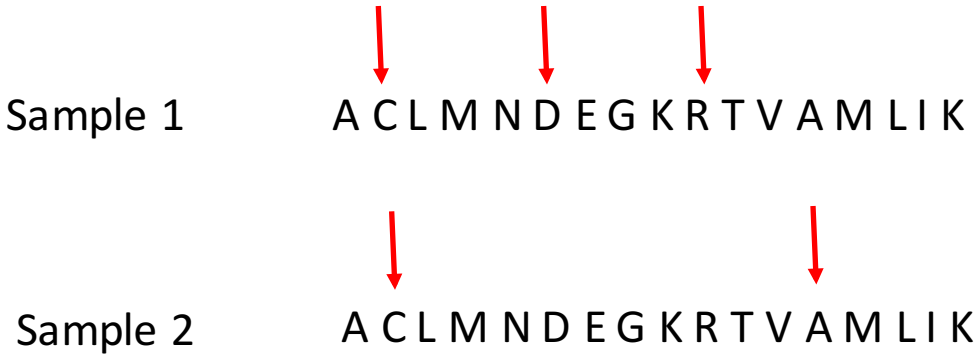
For the real "passenger" mutation , relative populations of the active and inactive states doesn't change even if they combine with additional passenger mutations. ($\Delta G_{\mathrm{wt}} \approx \Delta G_{\mathrm{pm}} \approx \Delta G_{\mathrm{pm}}$)

For the latent "driver" mutation , alters the relative populations of the active and inactive state but change is not sufficient on it's own. ($\Delta G_{\mathrm{wt}} < \Delta G_{\mathrm{lm}} << \Delta G_{\mathrm{DM}}$)

Nussinov et. al. 2015

Functional impact score based identification of latent drivers in coding regions

For a given gene G



FIS(C2), FIS(D6), FIS(R10) , FIS(A13)

FIS(C2 | C2,D2), FIS(C2 | C2,R10) , FIS(D6 | C2,R10)

FIS(C2| C2,D6,R10), FIS(D6 | C2,D6,R10), FIS(R10| C2,D6,R10)

FIS(C2| C2,A13), FIS(A13|C2,A13)

FIS(C2 | C2,D6)
FIS(C2 | C2,R10)
FIS(C2 | C2,D6,R10)
FIS(C2 | C2,A13)

$\chi(FIS(C2))$

χ -> Median / Min

True Passenger mutation

FIS (C2) ≈ $\chi(FIS(C2))$ << FIS(driver)

Latent Driver

FIS(C2) < $\chi(FIS(C2))$ << FIS(driver)

Other potential attributes of Latent Driver

Latent driver should have modestly elevated mutation frequency relative to background

Multiple latent driver will promote tumorigenesis through a polygenic model. Presence of genomic instability should facilitate the polygenic model.

They will be often linked to CNVs and large structural variations. Biochemical connections to known TSGs and oncogens.

Passenger mutation in pan-cancer dataset but latent driver in specific cancer type.