# How we analyze cancer?

El càncer és un grup de malalties caracteritzat per la proliferació descontrolada i desdiferenciació cel·lular, a més, té una evolució tipus darwiniana. Cada cèl·lula té una fitness i sobreviuen les que la tenen millor; les cèl·lules acumulen dany en el DNA (cèl·lules independents) i alguns d’aquests donen mutacions que aporten més capacitat de dividir-se (aquestes mutacions s’anomenen drivers).

El càncer és una successiva de mutacions driver que condueixen fins a la metàstasi. Els divers poden saber-se mitjançant la biologia del càncer i poden desenvolupar-se teràpies de prevenció, diagnòstic, tractament i prognosi.

Per veure les mutacions i variants pròpies de cada pacient es busquen els canvis entre una mostra de tumor d’aquest comparat amb un DNA de referència.

Els tumors tenen entre centenars i centenars de milers de mutacions; hi ha molta variabilitat fins i tot dins el mateix grup de càncer i poques d’aquestes mutacions seran mutacions driver, per saber quines poden ser-ho s’utilitzen mètodes computacionals.

* Seqüenciació de tumors cancerígens de diferents pacients i càncers
* Veure com es troben les mutacions somàtiques en el genoma

Un driver té més recurrència de mutacions que les esperades a l’atzar i tenen un efecte o impacte funcional, a més, tenen una tendència a acumular-se en una regió concreta de la seqüència (clustering).

Es desenvolupen diferents mètodes per trobar gens driver o regions no codificants driver, i també diferents senyals de selecció positiva ja que no tots els gens driver tenen les mateixes.

Atzar: acumulació de mutacions sense selecció positiva que dependrà de diferents factors com la compactació, temps de replicació, etc. És difícil introduir aquestes covariables, però les taxes de mutació també depenen de la seqüència mutada, dels nucleòtids que tenen upstream i downstream, cosa que permetrà simplificar-ho.

IntOGen: pipeline per trobar gens drivers.

Mitjançant dades de pacients amb càncer i mètodes computacionals es busquen gens drivers. Hi ha sis mètodes diferents: dos es basen en la recurrència, tres en el clustering i un amb l’impacte funcional.

Cada mètode proposa una lista de drives amb un p-value associat. S’han de combinar i segons com detecten els gens coneguts involucrats en càncer se’ls dona una valoració diferent als nous detectats. La combinació funciona millor que individualment, s0han trobat molts fens driver nous i, a més, alguns coneguts en un tipus de càncers s’han vist driver en un altre.

Cerca d’un gen, cohort o tipus de càncer en IntOGen.

OncodriveCLUSTL: com a imput se li donen les mutacions somàtiques de les cohorts. S’analitzen les SNV (single nucleotide variant).

1. Es mapeja la seqüència nucleotídica cercant-ne les mutacions.
2. Cerca de clústers: número de mutacions, separacions i score associat.
3. Suma dels scores de cada clúster donant un clustering observat. El clustering esperat es calcula obrint finestres per a cada mutació i es mira la probabilitat de mutar repetint-ho 1000 vegades mirant com s’agrupen. S’associa una score total
4. Comparar el clustering observat i l’esperat.

Per tenir la llista completa, és important utilitzar diferents mètodes que es basin en diferents senyals.

Cancer Genome Interpreter: cerca de mutacions driver, segons on es troben i si són o no drivers es mapegen. Les mutacions detectades com driver se cerquen en una base de dades per veure els tractaments que hi ha.