Bayes and NGS

JK

2019 9 4

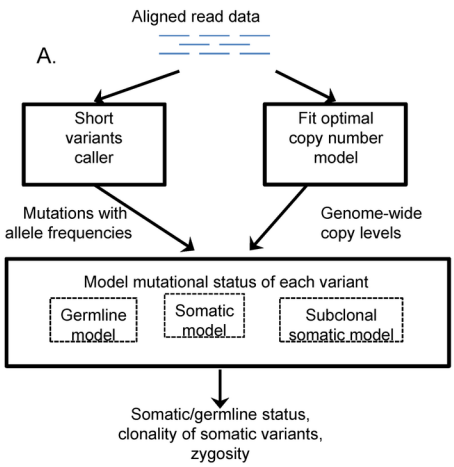
NGS process

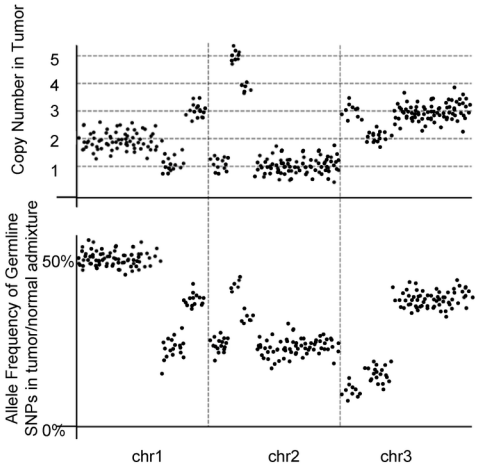
- Library preparation
  - DNA pool (diversity, many cells)
  - Targeted (multiplexing PCR, Hybrid capture)
  - Adeptor ligation (ID, primer, platform adaptor)
- Sequencing
  - Amplification (to enhance signal)
  - Sequencing (capture photo, check electric current change)
- Alignment
  - SAM, BAM file
- Variant calling
  - Bayes approach

$$\text{Posterior} = (\text{Prior} * \text{Likelyhood})/\text{Normalizing constant}$$

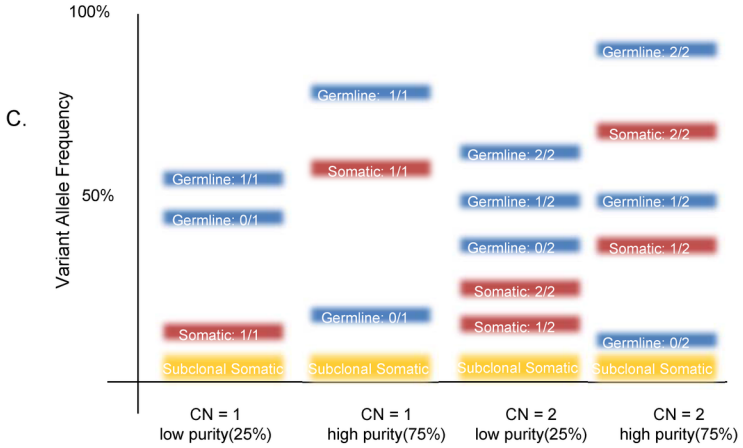
Allele frequency (AF)

- Allele
  - Population
  - Cancer
- Somatic vs Germline
  - CNV
  - Minor allele frequency
  - Tumor cellularity





copy number	LOH status	status of variant	0	5
C=1	LOH	V=0 germline	50	49
		V=1 somatic	0	3
		V=1 germline	50	51
C=2	LOH	V=0 germline	50	48
		V=2 somatic	0	5
		V=2 germline	50	53
	het	V=1 somatic	0	3
		V=1 germline	50	50
C=3	LOH	V=0 germline	50	46
		V=3 somatic	0	7
		V=3 germline	50	54
	het	V=1 somatic	0	2
		V=1 germline	50	49
		V=2 somatic	0	5
		V=2 germline	50	51
C=4	LOH	V=0 germline	50	45
		V=4 somatic	0	10
		V=4 germline	50	55
	het	V=1 somatic	0	2
		V=1 germline	50	48
		V=3 somatic	0	7
		V=3 germline	50	52
	het	V=2 somatic	0	5
		V=2 germline	50	50
C=5	LOH	V=0 germline	50	44
		V=5 somatic	0	12



- CNV
  - CNV
  - Minor allele frequency
  - Tumor cellularity
- Markov Chain Monte Carlo (MCMC)
  - Hierachical model
  - Determine parameter distribution by simulation