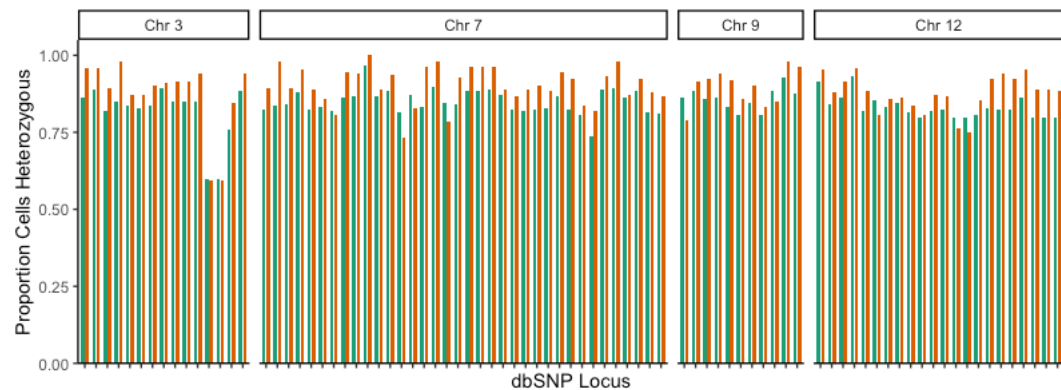
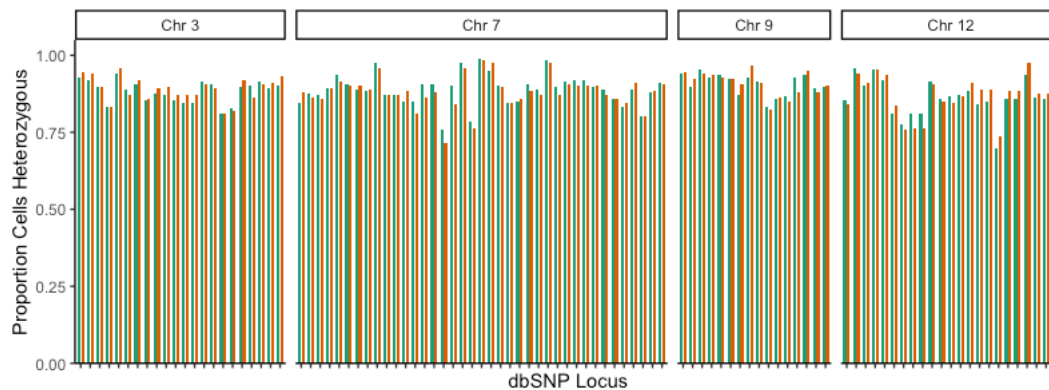


# Supplementary Information (Supp. Figures)

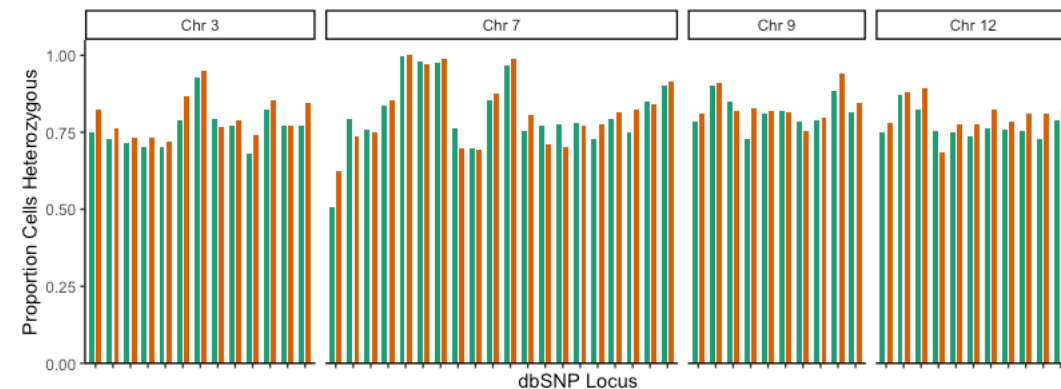
CCM 5005



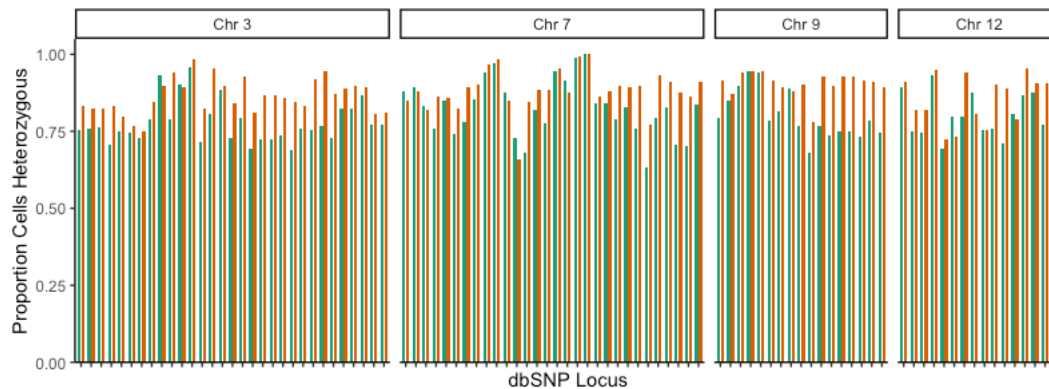
CCM 5006



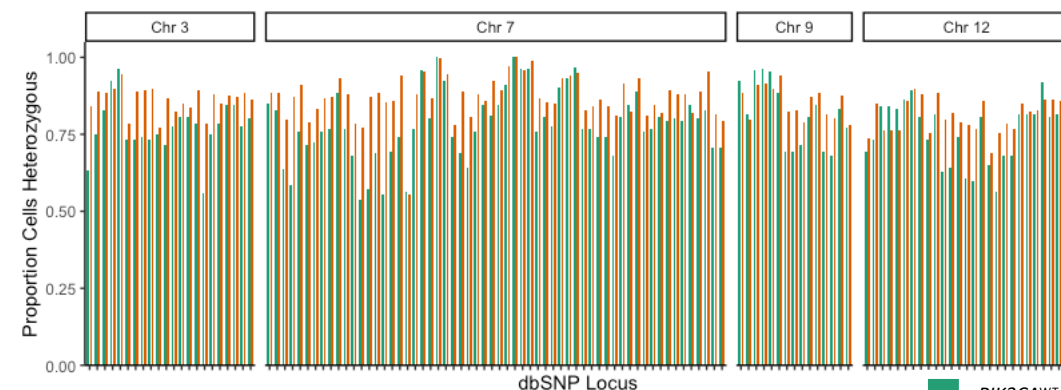
CCM 5035



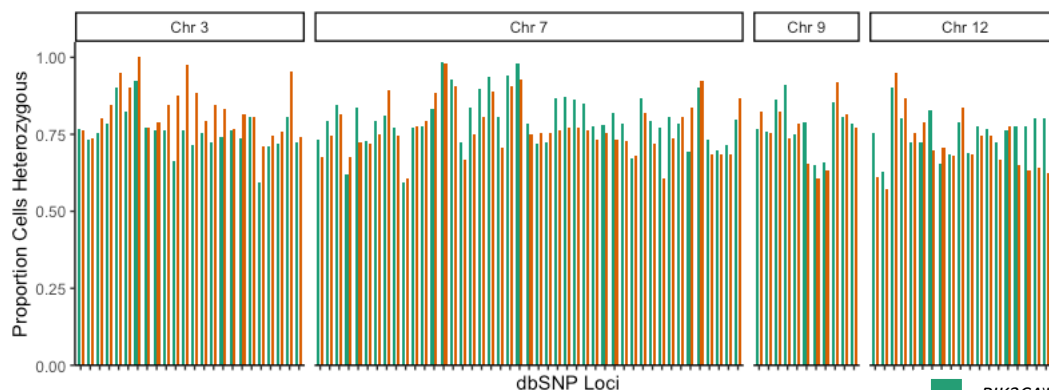
CCM 5009



CCM 5039



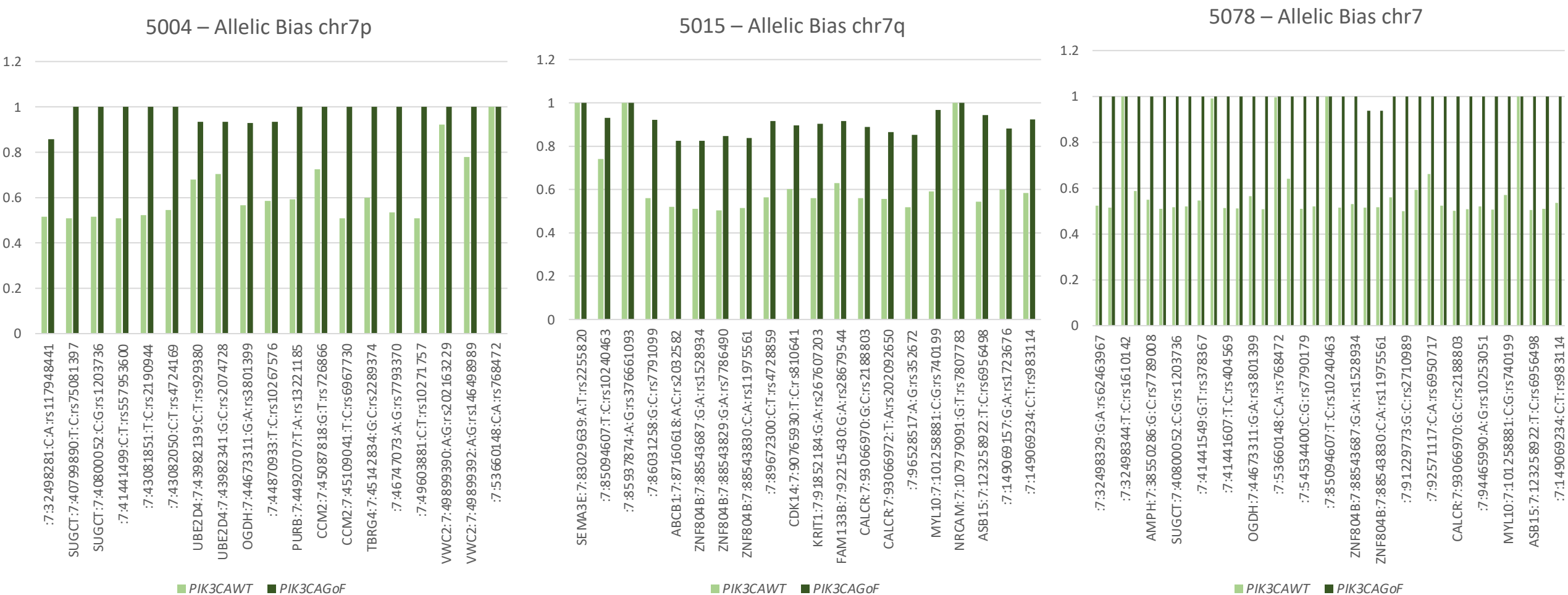
CCM Hu150



PIK3CA<sup>WT</sup>  
PIK3CA<sup>Gof</sup>

PIK3CA<sup>WT</sup>  
PIK3CA<sup>Gof</sup>

Supplementary figure 1: No evidence of somatic LOH in 6 CCM lesions. All graphs show the proportion of cells heterozygous for a given SNP for cells with (orange) and without (teal) an activating variant in *PIK3CA*. No set of 4 SNPs in a given bin or chromosome arm shows evidence of significant (FDR-corrected  $p < .05$ ) reduction in heterozygosity. Exact chromosomal location and rsids for all loci in graph are reported in Supplementary Data 5. Underlying data for bar charts in source data.

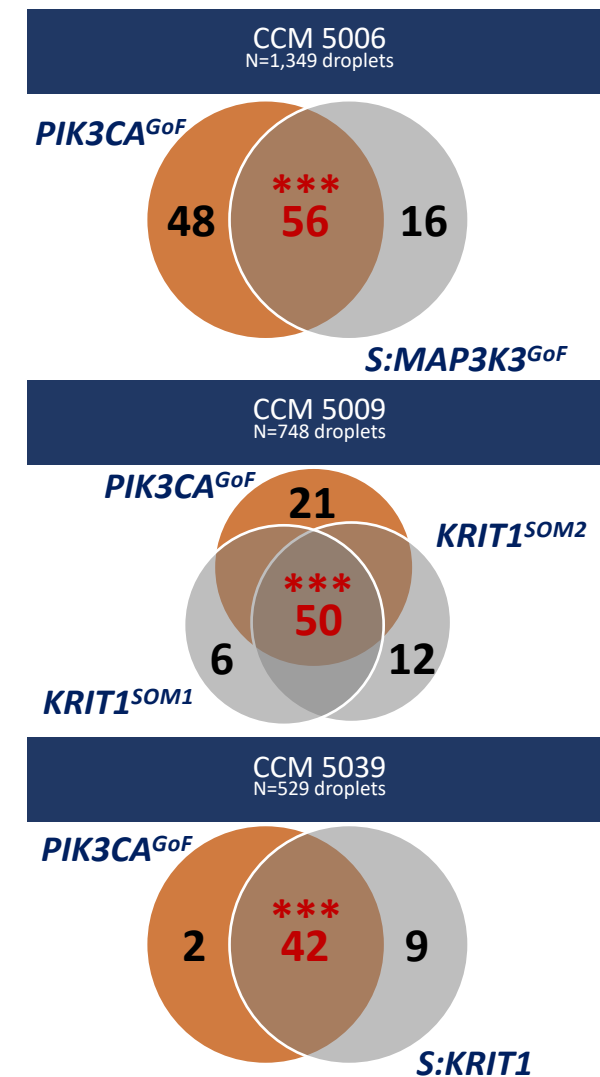


Supplementary figure 2: Allelic bias in regions of somatic LOH. Within the chromosomal arms with LOH identified in figure 3, we plotted the proportion of homozygous calls pertaining to the predominant allele in cells with (dark green) and without (light green) an activating variant in *PIK3CA*. A score of 0.5 is suggestive of random allelic dropout, with each parental allele called as homozygous in equal proportions, whereas a score of 1 means all homozygous calls at a given SNP are for a single parental allele. Underlying data for bar charts included in source data for figure 2.

a

Sample ID	Inheritance	PIK3CA variant	Lesion Variant 1	Lesion Variant 2
5004	Sporadic	Yes (1%)	S: <i>CCM2</i>	LOH (chr7p)
5005	Sporadic	Yes (2%)	?	?
5006	Sporadic	Yes (8%)	S: <i>MAP3K3</i>	
5009	Sporadic	Yes (<1%)	S: <i>KRIT1</i>	S: <i>KRIT1</i>
5015	Familial	Yes (4.2%)	G: <i>KRIT1</i>	LOH (chr7q)
5035	Sporadic	Yes (4.2%)	?	?
5039	Familial	Yes (9%)	G: <i>KRIT1</i>	S: <i>KRIT1</i>
5078	Familial	Yes (1%)	G: <i>CCM2</i>	LOH (chr7)
Hu150	Sporadic	Yes (11%)	?	?

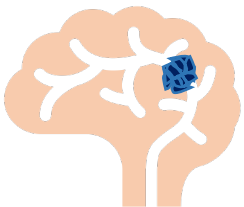
b



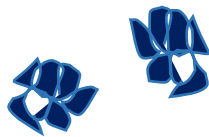
Supplementary figure 3: Summary of genetic diagnoses in CCMs. a) We identified somatic LOH or previously unidentified variants in 6/9 lesions (in gray). b) The identified variants are enriched in cells with activating variants in *PIK3CA* for CCMs 5006, 5009 and 5039. S: Somatic, G: Germline. All variants are specifically reported in Supplementary Data 6. \*\*\*  $p < 1e-4$ . \*  $p < .01$ . Exact p-values and underlying data for (b) are available in source data, including source data for figures 2a and 3c,d

a

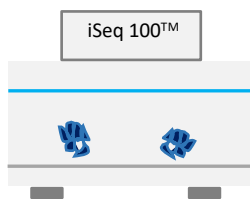
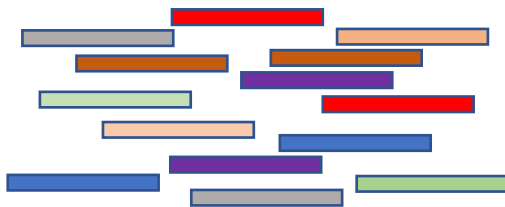
Surgical Resected Lesion 5009



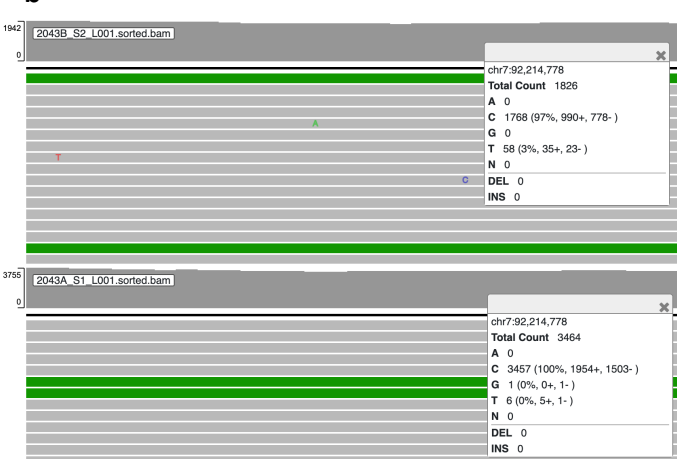
Sampled from discrete parts of lesion in OCT block



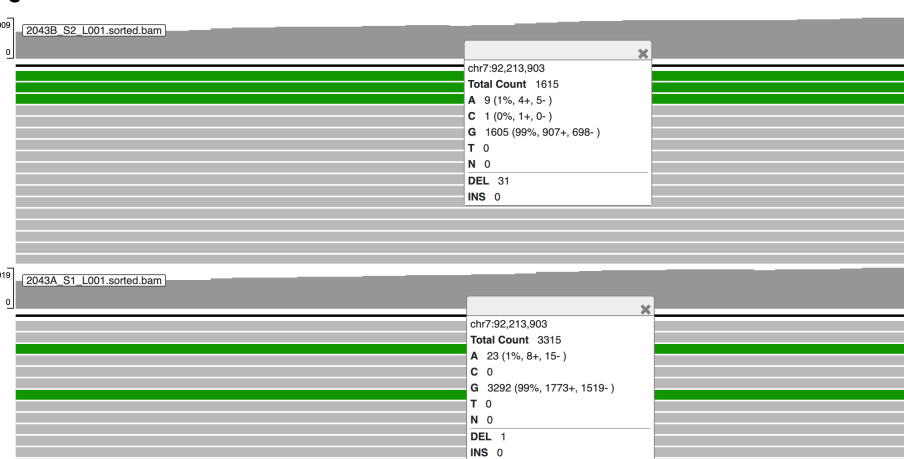
Libraries generated using targeted gene panel including CCM genes and oncogenes and sequenced with iSeq



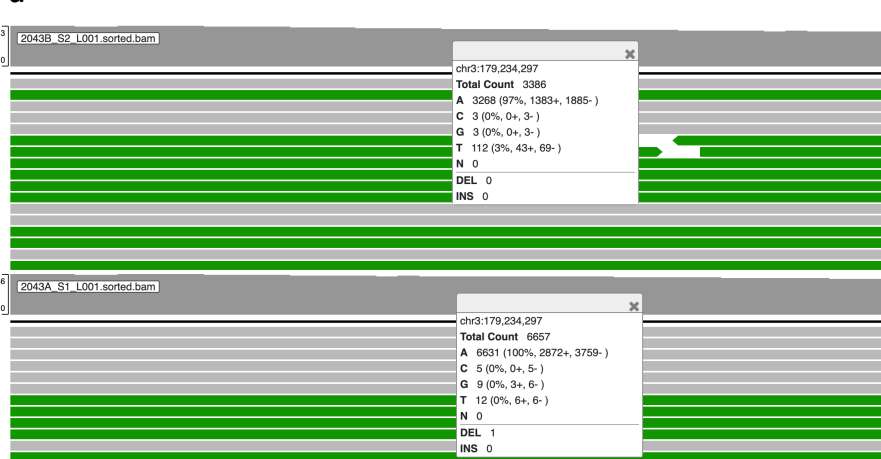
b



c

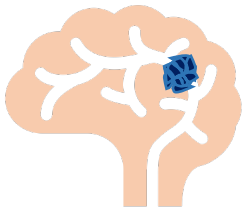


d

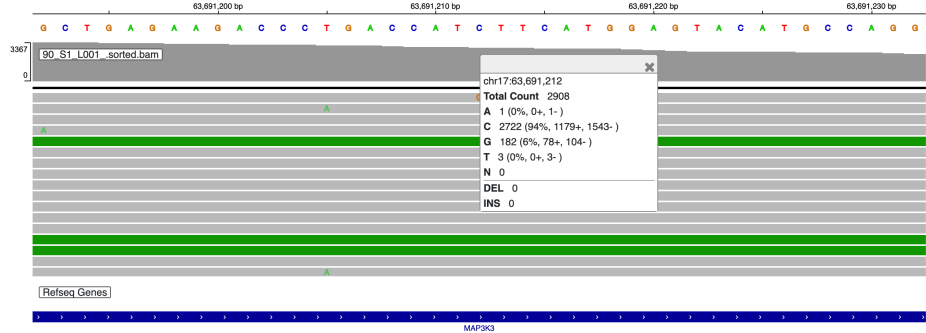
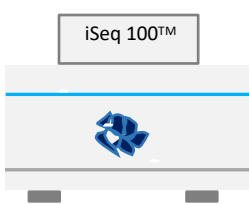


e

Surgical Resected Lesion 5006



Sequenced remaining lesion with iSeq



Supplementary figure 4: Genetic heterogeneity in CCM 5009. a) Schematic showing targeted bulk DNA-sequencing of two sections of resected lesion CCM 5009. All three variants from CCM lesion 5009 that were identified using snDNA-sequencing (b) hg38 chr7:92214778:C/T, (c) hg38 chr7:92213903:GA/A & (d) hg38 chr3:179234297:A/T) were similarly identified in ~1.8 - 3.3% of reads in one section of lesion 5009 (2043B, (b),(c),(d) top). However, all three variant sequences were identified in less than 0.2% of reads in a second subsection of the same lesion (2043B, (b),(c),(d) bottom) processed, sequenced and analyzed simultaneously, suggesting heterogeneity within the lesion can confound identification of pathogenic variants. e) For CCM 5006, we validated the *MAP3K3* variant hg38 chr17:63691212:C>G with similar strategy, using same panel and targeted bulk DNA-sequencing on one alternative section of the lesion.