Supplemental data

Table S1 Common regions between two panels (ghrch38)

chr3	179203735	179203784	PIK3CA
chr3	179210162	179210198	PIK3CA
chr3	179210278	179210302	PIK3CA
chr3	179218278	179218320	PIK3CA
chr3	179221106	179221154	PIK3CA
chr3	179234277	179234314	PIK3CA
chr6	152011685	152011744	ESR1
chr6	152094382	152094421	ESR1
chr6	152098779	152098822	ESR1
chr7	55191750	55191833	EGFR
chr12	25245337	25245376	KRAS
chr12	56085039	56085087	ERBB3
chr12	56088129	56088171	ERBB3
chr12	56088530	56088578	ERBB3
chr12	56088808	56088860	ERBB3
chr12	56098838	56098900	ERBB3
chr14	104780185	104780228	AKT1

chr17	7669648	7669693	TP53
chr17	7670567	7670643	TP53
chr17	7670692	7670727	TP53
chr17	7673532	7673609	TP53
chr17	7673686	7673722	TP53
chr17	7673760	7673839	TP53
chr17	7674187	7674264	TP53
chr17	7674863	7674896	TP53
chr17	7674940	7674972	TP53
chr17	7675061	7675099	TP53
chr17	7675138	7675219	TP53
chr17	7675962	7676008	TP53
chr17	7676054	7676127	TP53
chr17	7676174	7676208	TP53
chr17	7676253	7676297	TP53
chr17	7676376	7676419	TP53
chr17	7676435	7676440	TP53
chr17	7676545	7676612	TP53
chr17	39723961	39723993	ERBB2

Table S2: Overview of the detected variants and VAF by the Oncomine and the Qiaseq panel

Subject	Gene	Variant	VAF Oncomine	VAF Qiaseq
55				
68	PIK3CA	p.E545K	1.8	x
	PIK3CA	p.G1049R	2.6	x
	TP53	p.R280K	1.0	x
	TP53	p.P250L	0.4	x
79				
85	PIK3CA	p.E545K	49.2	45.2
	CDH1	p.Q23*	NA	26.2
97	PIK3CA	p.H1047R	0.2	x
145	PTEN	p.D92E	NA	18.7
	CDH1	c.1565+1G>A	NA	17.1
147				
159	PIK3CA	p.H1047R	0.3	x
172	AKT1	p.E17K	35.5	32.2
	CBFB	p.Q67H	NA	26.1
211				
219				
226	PIK3CA	p.E545K	26.5	24.8
237	PIK3CA	p.E453K	13.85	6.5
	PIK3CA	p.H1047R	8.9	17.2
253	PIK3CA	p.H1047R	8.5	x
286				
296				
304				
306	PIK3CA	p.H1047R	14.9	18.6
	PIK3CA	p.P539R	x	11.5
	KRAS	p.G12V	1.1	x
316	PIK3CA	p.H1047R	1.0	x
317	KMT2C	p.K822R	NA	1.3
335				
341	PIK3CA	p.E542K	1.0	3.2
350	KMT2C	p.H367Y	NA	1.5
	TP53	p.K132R	1.2	x
357				
361				
371				
376	ESR1	p.D538G	5.7	10.4
376	ESR1	p.E380Q	2.0	x
381	PIK3CA	p.E545K	29.9	32.5

	TP53	c.673-1G>T	NA	31.4
391	TP53	p.R273C	53.3	42.3
397				
407				
412				
415				
422	AKT1	p.E17K	12.1	17.4
427	PIK3CA	p.E545K	3.6	x
	PIK3CA	p.E726K	3.1	5.2
435				
448				
460	TP53	tvc.novel.2	0.9	x
	PIK3CA	p.G1049R	3.1	x
	PIK3CA	p.Q546K	3.2	3.6
467	PIK3CA	p.E545K	3.8	x
	PIK3CA	p.E726K	2.3	x
	TP53	p.R280K	6.1	x
468				
471	PIK3CA	p.E726K	0.4	x
	PIK3CA	p.N345K	22.9	15.3
475	PIK3CA	p.N345K	2.0	x
476	PIK3CA	p.H1047R	12.4	15.4
	SMARCA4	p.F1142L	NA	19.1
479	PIK3CA	p.E542K	26.2	30.8
	CDH1	p.Q23*	NA	10.5
481	ERBB2	p.L755S	5.3	x
	PIK3CA	p.H1047L	7.5	12.2
484	PIK3CA	p.H1047R	2.2	x
486	PIK3CA	p.N345K	23.2	31.3
486	TP53	p.R282W	0.9	x
491	PIK3CA	p.E545K	11.0	26.7
492				
493				
500				
513	ERBB3	p.D297Y	1.8	x
519	PIK3CA	p.E545K	7.3	5.8
521				
524	PIK3CA	p.E545K	0.9	x
529	PIK3CA	p.H1047R	1.6	x
535				
541				

570				
592	PIK3CA	p.H1047R	3.0	5.4
600				
602	TP53	p.G199V	NA	5.1
609	PIK3CA	p.H1047R	0.7	x
618				
638				
659	TP53	p.H193R	3.2	x
661	PIK3CA	p.E542K	46.7	40.9
661	TP53	p.P151H	15.7	15.9
661	PIK3CA	p.H1047R	4.4	x
670	TP53	p.G244D	36.2	36.0
680	PIK3CA	p.H1047R	1.4	x
692	PIK3CA	p.H1047R	5.0	X
692	CDH1	c. 1565 +1G>A	NA	7.1

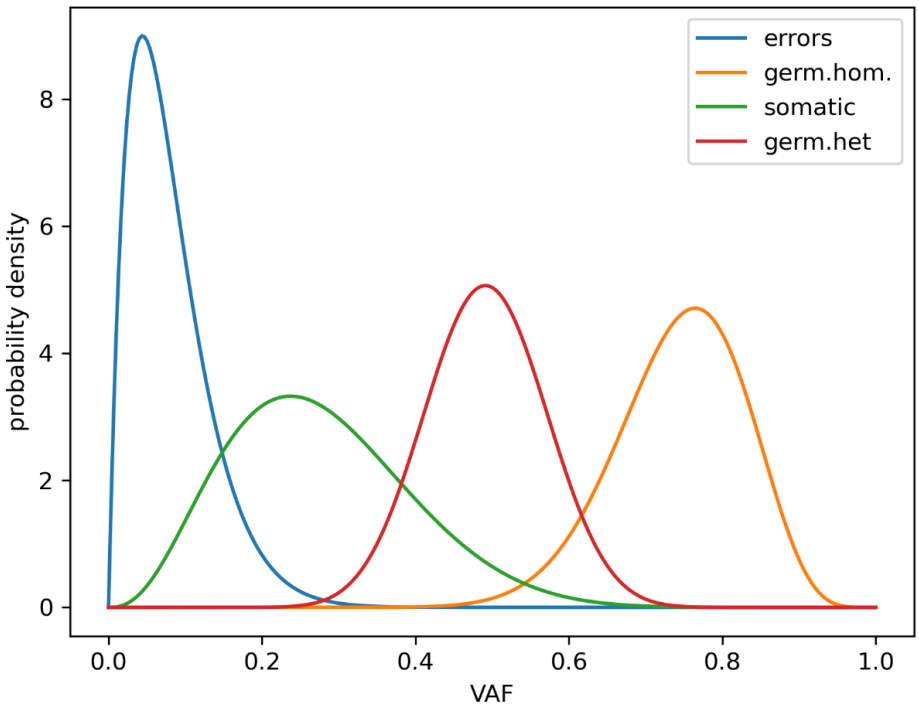


Figure S1: Beta distributions of the four expected types of variants. Errors: sequencing errors, germ.hom: germline heterozygous variants, somatic: somatic variants, germ.het: germline homozygous variants.

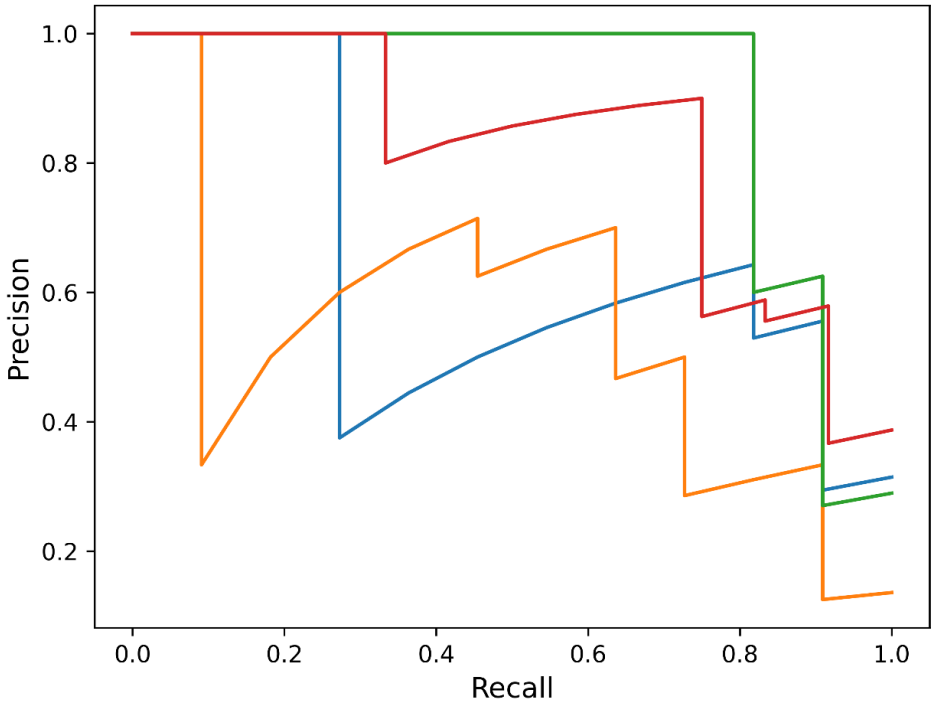


Figure S2 Precision-recall curves of our SVM model using 4-fold cross validation.

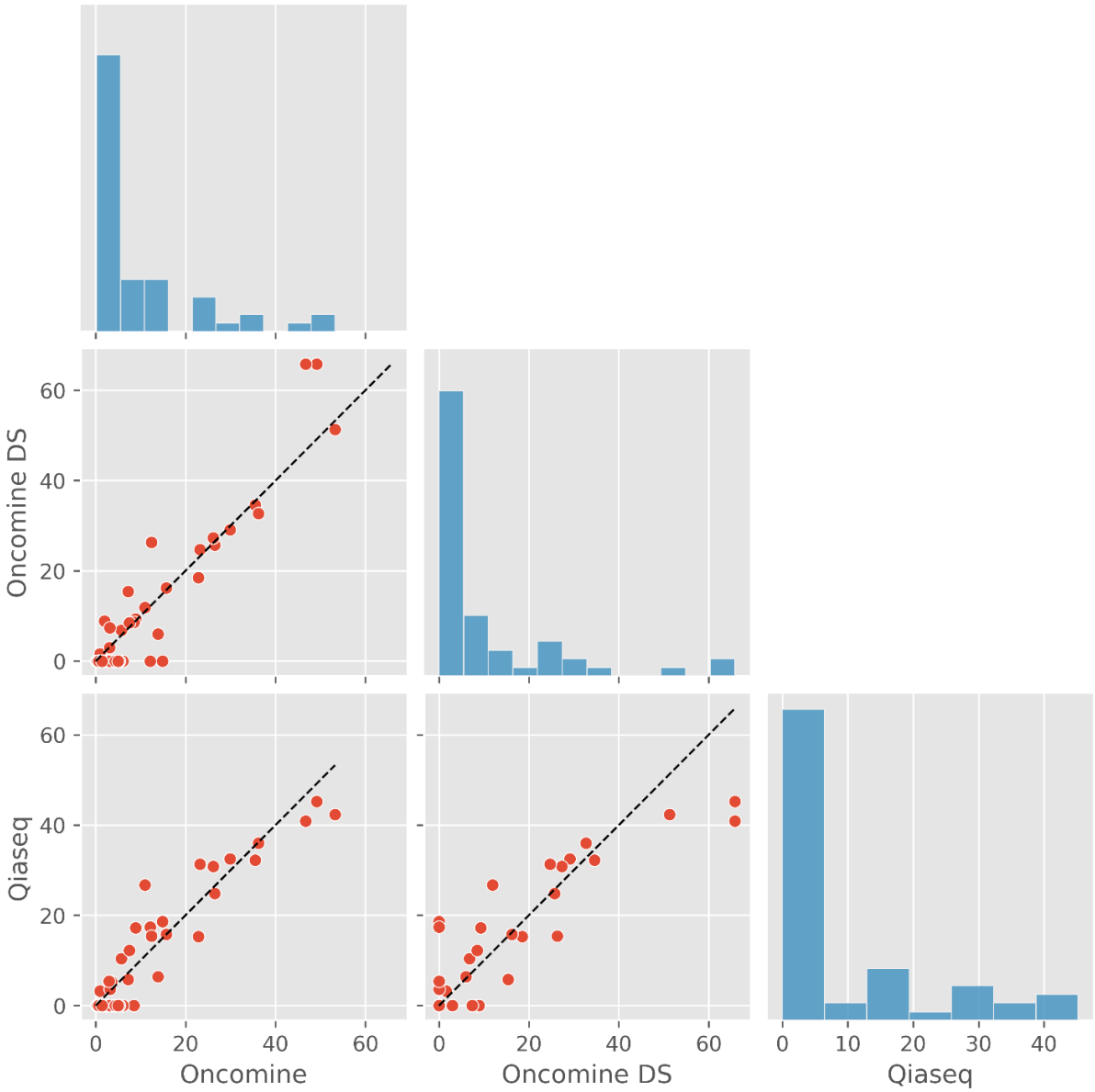
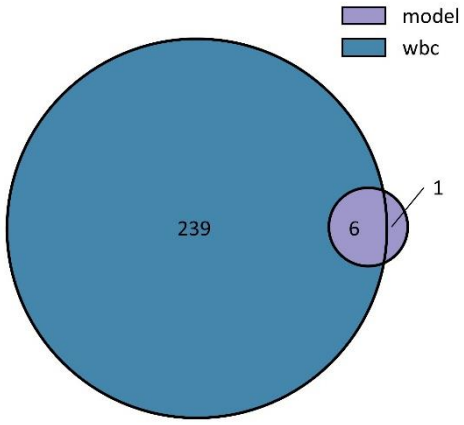


Figure S3: Comparison of the VAFs for mutations detected in the common regions between Qiaseq and Oncomine. Plots along the diagonal show the VAF distributions for Oncomine, down-sampled Oncomine (Oncomine DS) and Qiaseq. In off-diagonal plots, each red dot corresponds to a mutation and its *x* and *y* co-ordinates represent the VAF of that mutation in the corresponding panel. The dashed black line show the *y=x* line. For mutations that were not detected we set the VAF to zero.

A.



B.

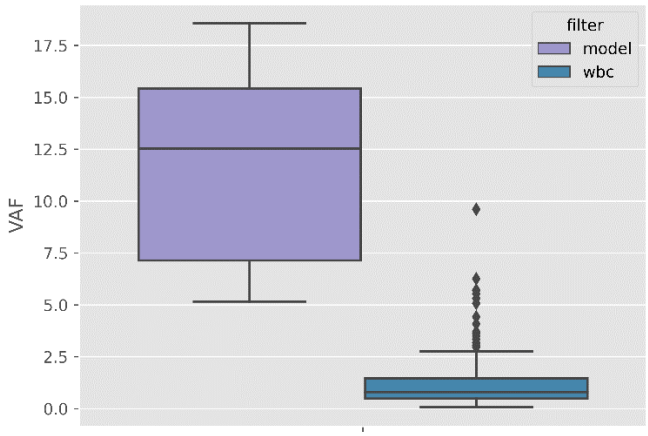


Figure S4: Comparison of model-based filtering of Qiaseq variant calls to the use of matched germline data from white blood cells. (A) Venn diagram showing the numbers of variants detected using the two approaches. (B) Comparison of the VAF distributions of those variants.