

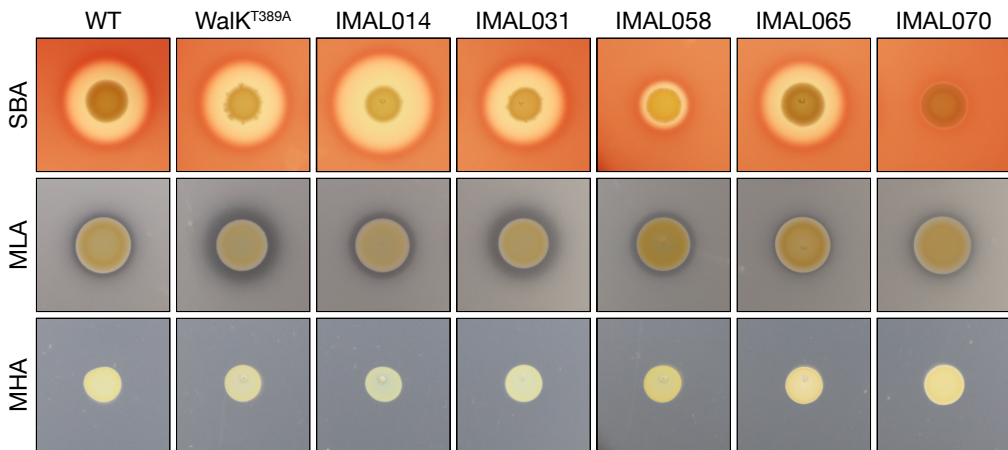


Figure S1: Phenotypic characterisation of the isolates.

SBA = Sheep blood agar (5% sheep blood in Columbia agar); shows haemolytic activity.

MLA = *Micrococcus luteus* agar (BHI agar containing 0.2% *Micrococcus luteus* cells [Sigma-Aldrich, M3770]); shows the secretion/activity of one WalR-regulated peptidoglycan hydrolase (Atl).

MHA = Mueller-Hinton agar (Difco); shows the colour of the cells. IMAL058 is a SigB mutant, making it more yellow than the other isolates.

The figure consists of two bar charts. The top chart displays 'SNP errors' on a logarithmic y-axis (0 to 1000). The bottom chart displays 'Indel errors' on a logarithmic y-axis (0 to 100). The x-axis for both charts is divided into four categories: Read-based, Short-read assemblies, Long-read assemblies, and Hybrid assemblies. Each category contains multiple bars representing different sequencing methods and coverage levels (20x, 50x, 100x). The bars are colored blue, except for the first two bars in the Read-based category which are orange. Above each bar, there are two rows of numerical values: 'SNP sens' and 'SNP prec' for the top chart, and 'indel sens' and 'indel prec' for the bottom chart.

Category	Method	Coverage	SNP sens	SNP prec	indel sens	indel prec	
Read-based	Illumina	20x	1.0	1.0	1.0	1.0	
		50x	1.0	1.0	1.0	1.0	
	ONT	20x	0.78	0.96	1.0	1.0	
		50x	1.0	1.0	1.0	1.0	
		100x	1.0	1.0	1.0	1.0	
		100x	1.0	1.0	1.0	1.0	
Short-read assemblies	Shovill	20x	1.0	0.013	1.0	0.13	
		50x	1.0	0.35	1.0	0.20	
		100x	1.0	0.37	1.0	0.22	
	Unicycler	20x	1.0	0.057	1.0	0.29	
		50x	1.0	0.45	1.0	1.0	
		100x	1.0	0.74	1.0	1.0	
	SKESA	20x	1.0	0.25	1.0	0.33	
		50x	1.0	0.66	1.0	1.0	
		100x	1.0	0.88	1.0	1.0	
	Long-read assemblies	Canu	20x	1.0	0.66	1.0	0.10
			50x	1.0	0.92	1.0	0.29
			100x	1.0	0.96	1.0	1.0
Flye		20x	1.0	0.85	1.0	0.18	
		50x	1.0	1.0	1.0	0.67	
		100x	1.0	0.68	1.0	0.080	
Raven		20x	1.0	0.39	1.0	0.023	
		50x	1.0	0.72	1.0	0.057	
		100x	1.0	0.88	1.0	0.071	
Hybrid assemblies	Unicycler	20x	1.0	0.055	1.0	0.080	
		50x	1.0	0.10	1.0	0.40	
		100x	1.0	0.14	1.0	0.50	
	Hybracter	20x	1.0	1.0	1.0	1.0	
		50x	1.0	1.0	1.0	1.0	
		100x	1.0	1.0	1.0	1.0	

Figure S2: Variant calling metrics for both read- and assembly-based variant calling methods at each read depth. False negative and false positive errors are shown in the plots. Sensitivity (sens) and precision (prec) are shown above the plots. The y-axes have a pseudo-log transformation.

Variant-calling errors by long-read polishing

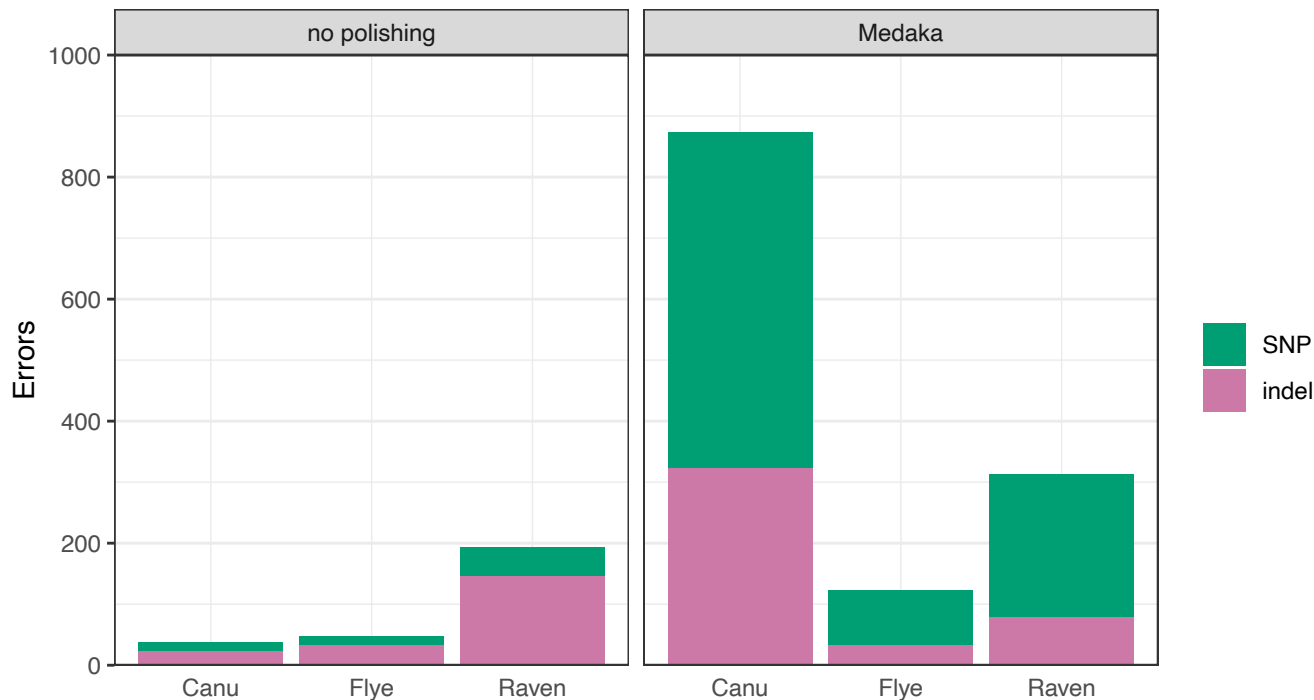


Figure S3: Variant-calling errors, before and after Medaka polishing, for each of the long-read assembly methods.