Pipeline Report

J. Matthew Holt

May 19, 2022

Contents

1 Metadata information

1.1 Constants

The following information is not automatically updated in the report:
Build: b38 (hg38)

1.2 Dynamic

The following information is gathered dynamically when a report is generated:

Library	Sample	Preparation	Gbp
HALB3002753	NA12878	PCR-free-illumina	181.14
HALB3002754	HG002	PCR-free-illumina	145.87
HALB3002755	HG003	PCR-free-illumina	149.77
HALB3002756	HG004	PCR-free-illumina	138.30
HALB3002757	HG005	PCR-free-illumina	159.85
HALB3003200	NA12878	PCR-free-illumina	167.53
HALB3003201	HG002	PCR-free-illumina	129.81
HALB3003222	HG003	PCR-free-illumina	156.20
HALB3010451	NA12878	PacBio-CCS-high	77.46
HALB3010452	HG002	PacBio-CCS-high	166.22
HALB3010453	HG003	PacBio-CCS-high	129.52
HALB3010454	HG004	PacBio-CCS-high	123.81
HALB3010455	HG005	PacBio-CCS-high	145.34
HALB3010456	HG006	PacBio-CCS-high	122.17
HALB3010457	HG007	PacBio-CCS-high	113.30
HALB3012850	HG006	PCR-free-illumina	127.71
HALB3012851	HG007	PCR-free-illumina	127.67

2 Summarized Small Variant Comparison

This section contains the summarized results of the small variant analysis. For each aligner-caller pair, we ran RTG on the Genome in a Bottle (GIAB) high confidence truth set ($\approx$3.6M variants). The following results are gathered from the RTG 'summary.txt' file using the final row (i.e. "None" indicating no variants were filtered) and with the `--all-records` option enabled.

2.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
clinical	sentieon-201808.07	strelka-2.9.10	10	0.9920+-0.0014	0.9918+-0.0008	0.9919+-0.0010
hg38_T2T_masked	dragmap-1.2.1	clair3-0.1-r11	10	0.9975+-0.0003	0.9920+-0.0006	0.9947+-0.0003
hg38_T2T_masked	dragmap-1.2.1	dnascope-1.0-202112.01-PO	10	0.9984+-0.0003	0.9941+-0.0009	0.9963+-0.0005
hg38_T2T_masked	dragmap-1.2.1	octopus-0.7.4	10	0.9961+-0.0003	0.9897+-0.0008	0.9929+-0.0005
hg38_T2T_masked	dragmap-1.2.1	strelka-2.9.10	10	0.9889+-0.0036	0.9920+-0.0009	0.9904+-0.0021
hg38_T2T_masked	sentieon-202112.01	clair3-0.1-r11	10	0.9972+-0.0004	0.9927+-0.0008	0.9950+-0.0005
hg38_T2T_masked	sentieon-202112.01	dnascope-1.0-202112.01-PO	10	0.9983+-0.0003	0.9943+-0.0008	0.9963+-0.0005
hg38_T2T_masked	sentieon-202112.01	octopus-0.7.4	10	0.9945+-0.0003	0.9901+-0.0008	0.9923+-0.0005
hg38_T2T_masked	sentieon-202112.01	strelka-2.9.10	10	0.9864+-0.0037	0.9930+-0.0007	0.9897+-0.0021
hg38_T2T_masked	sentieon-202112.01-recal	clair3-0.1-r11	10	0.9973+-0.0003	0.9918+-0.0009	0.9945+-0.0005
hg38_T2T_masked	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	10	0.9985+-0.0002	0.9935+-0.0009	0.9960+-0.0005
hg38_T2T_masked	sentieon-202112.01-recal	octopus-0.7.4	10	0.9946+-0.0003	0.9901+-0.0008	0.9923+-0.0005
hg38_T2T_masked	sentieon-202112.01-recal	strelka-2.9.10	10	0.9903+-0.0015	0.9929+-0.0007	0.9916+-0.0011
hg38_T2T_masked	snap-2.0.0	clair3-0.1-r11	10	0.9964+-0.0004	0.9921+-0.0010	0.9942+-0.0006
hg38_T2T_masked	snap-2.0.0	dnascope-1.0-202112.01-PO	10	0.9976+-0.0003	0.9942+-0.0007	0.9959+-0.0005
hg38_T2T_masked	snap-2.0.0	octopus-0.7.4	10	0.9937+-0.0003	0.9905+-0.0007	0.9921+-0.0005
hg38_T2T_masked	snap-2.0.0	strelka-2.9.10	10	0.9855+-0.0039	0.9933+-0.0007	0.9894+-0.0023
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r11	10	0.9975+-0.0002	0.9698+-0.0012	0.9834+-0.0006
hg38_asm5_alt	dragmap-1.2.1	dnascope-1.0-202112.01-PO	10	0.9978+-0.0002	0.9737+-0.0009	0.9856+-0.0006
hg38_asm5_alt	dragmap-1.2.1	octopus-0.7.4	10	0.9968+-0.0002	0.9681+-0.0010	0.9822+-0.0005
hg38_asm5_alt	dragmap-1.2.1	strelka-2.9.10	10	0.9897+-0.0036	0.9699+-0.0010	0.9797+-0.0022
hg38_asm5_alt	sentieon-202112.01	clair3-0.1-r11	10	0.9976+-0.0003	0.9921+-0.0009	0.9948+-0.0005
hg38_asm5_alt	sentieon-202112.01	dnascope-1.0-202112.01-PO	10	0.9984+-0.0002	0.9938+-0.0007	0.9961+-0.0004
hg38_asm5_alt	sentieon-202112.01	octopus-0.7.4	10	0.9957+-0.0003	0.9891+-0.0008	0.9924+-0.0005
hg38_asm5_alt	sentieon-202112.01	strelka-2.9.10	10	0.9885+-0.0036	0.9920+-0.0008	0.9902+-0.0021

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_asm5_alt	sentieon-202112.01-recal	clair3-0.1-r11	10	0.9977+-0.0003	0.9910+-0.0007	0.9943+-0.0004
hg38_asm5_alt	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	10	0.9986+-0.0002	0.9930+-0.0008	0.9957+-0.0005
hg38_asm5_alt	sentieon-202112.01-recal	octopus-0.7.4	10	0.9959+-0.0003	0.9891+-0.0008	0.9925+-0.0005
hg38_asm5_alt	sentieon-202112.01-recal	strelka-2.9.10	10	0.9920+-0.0014	0.9918+-0.0008	0.9919+-0.0010
hg38_asm5_alt	snap-2.0.0	clair3-0.1-r11	10	0.9966+-0.0004	0.9920+-0.0008	0.9943+-0.0005
hg38_asm5_alt	snap-2.0.0	dnascope-1.0-202112.01-PO	10	0.9977+-0.0003	0.9941+-0.0007	0.9959+-0.0004
hg38_asm5_alt	snap-2.0.0	octopus-0.7.4	10	0.9943+-0.0003	0.9903+-0.0007	0.9923+-0.0005
hg38_asm5_alt	snap-2.0.0	strelka-2.9.10	10	0.9862+-0.0039	0.9931+-0.0007	0.9896+-0.0023

2.2   Preparation: "PacBio-CCS-high"

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_T2T_masked	minimap2-2.23	clair3-0.1-r11	7	0.9980+-0.0013	0.9987+-0.0006	0.9983+-0.0010
hg38_T2T_masked	minimap2-2.23	dnascope_hifi-0.4-202112.01	7	0.9986+-0.0008	0.9986+-0.0006	0.9986+-0.0007
hg38_T2T_masked	minimap2-2.23	peppermargindv-r0.8	7	0.9983+-0.0009	0.9985+-0.0005	0.9984+-0.0007
hg38_T2T_masked	pbmm2-1.7.0	clair3-0.1-r11	7	0.9984+-0.0013	0.9987+-0.0006	0.9986+-0.0009
hg38_T2T_masked	pbmm2-1.7.0	dnascope_hifi-0.4-202112.01	7	0.9987+-0.0008	0.9987+-0.0005	0.9987+-0.0007
hg38_T2T_masked	pbmm2-1.7.0	peppermargindv-r0.8	7	0.9985+-0.0009	0.9984+-0.0005	0.9985+-0.0007
hg38_T2T_masked	sentieon_mm2-202112.01	clair3-0.1-r11	7	0.9980+-0.0013	0.9987+-0.0006	0.9983+-0.0010
hg38_T2T_masked	sentieon_mm2-202112.01	dnascope_hifi-0.4-202112.01	7	0.9986+-0.0008	0.9986+-0.0006	0.9986+-0.0007
hg38_T2T_masked	sentieon_mm2-202112.01	peppermargindv-r0.8	7	0.9983+-0.0009	0.9985+-0.0005	0.9984+-0.0007

3 Summarized SNP/Indel Comparison

This section contains the summarized results of the small variant analysis when stratified by SNP/Indel. Files were filtered using `bcftools filter` to pair down to SNPs or Indels only. For each aligner-caller pair, we ran RTG on the Genome in a Bottle (GIAB) high confidence truth set ($\approx$3.6M variants). The following results are gathered from the RTG 'summary.txt' file using the final row (i.e. "None" indicating no variants were filtered) and with the `--all-records` option enabled.

3.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
clinical	sentieon-201808.07	strelka-2.9.10	SNP	10	0.9909+-0.0015	0.9920+-0.0007	0.9915+-0.0010
			Indel	10	0.9888+-0.0010	0.9844+-0.0019	0.9866+-0.0015
hg38_T2T_masked	dragmap-1.2.1	clair3-0.1-r11	SNP	10	0.9975+-0.0003	0.9920+-0.0006	0.9948+-0.0003
			Indel	10	0.9949+-0.0006	0.9892+-0.0010	0.9921+-0.0007
hg38_T2T_masked	dragmap-1.2.1	dnascope-1.0-202112.01-PO	SNP	10	0.9986+-0.0002	0.9939+-0.0009	0.9962+-0.0005
			Indel	10	0.9968+-0.0005	0.9946+-0.0008	0.9957+-0.0006
hg38_T2T_masked	dragmap-1.2.1	octopus-0.7.4	SNP	10	0.9960+-0.0003	0.9851+-0.0011	0.9906+-0.0006
			Indel	10	0.9490+-0.0037	0.9631+-0.0044	0.9560+-0.0040
hg38_T2T_masked	dragmap-1.2.1	strelka-2.9.10	SNP	10	0.9873+-0.0041	0.9922+-0.0009	0.9897+-0.0024
			Indel	10	0.9886+-0.0011	0.9846+-0.0020	0.9866+-0.0015
hg38_T2T_masked	sentieon-202112.01	clair3-0.1-r11	SNP	10	0.9972+-0.0004	0.9929+-0.0008	0.9950+-0.0005
			Indel	10	0.9948+-0.0006	0.9892+-0.0015	0.9920+-0.0010
hg38_T2T_masked	sentieon-202112.01	dnascope-1.0-202112.01-PO	SNP	10	0.9985+-0.0002	0.9941+-0.0008	0.9963+-0.0005
			Indel	10	0.9968+-0.0005	0.9946+-0.0008	0.9957+-0.0006
hg38_T2T_masked	sentieon-202112.01	octopus-0.7.4	SNP	10	0.9943+-0.0003	0.9855+-0.0010	0.9899+-0.0006
			Indel	10	0.9483+-0.0037	0.9631+-0.0044	0.9557+-0.0039
hg38_T2T_masked	sentieon-202112.01	strelka-2.9.10	SNP	10	0.9847+-0.0042	0.9933+-0.0007	0.9890+-0.0023
			Indel	10	0.9873+-0.0011	0.9849+-0.0020	0.9861+-0.0015
hg38_T2T_masked	sentieon-202112.01-recal	clair3-0.1-r11	SNP	10	0.9974+-0.0003	0.9919+-0.0008	0.9946+-0.0005
			Indel	10	0.9942+-0.0007	0.9883+-0.0015	0.9913+-0.0010
hg38_T2T_masked	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	SNP	10	0.9987+-0.0002	0.9935+-0.0009	0.9961+-0.0005
			Indel	10	0.9966+-0.0006	0.9928+-0.0013	0.9947+-0.0009
hg38_T2T_masked	sentieon-202112.01-recal	octopus-0.7.4	SNP	10	0.9944+-0.0003	0.9854+-0.0010	0.9899+-0.0006
			Indel	10	0.9490+-0.0037	0.9632+-0.0044	0.9561+-0.0039

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_T2T_masked	sentieon-202112.01-recal	strelka-2.9.10	SNP	10	0.9891+-0.0016	0.9931+-0.0007	0.9911+-0.0011
			Indel	10	0.9873+-0.0011	0.9850+-0.0020	0.9861+-0.0015
hg38_T2T_masked	snap-2.0.0	clair3-0.1-r11	SNP	10	0.9963+-0.0004	0.9921+-0.0010	0.9942+-0.0006
			Indel	10	0.9942+-0.0006	0.9897+-0.0017	0.9919+-0.0011
hg38_T2T_masked	snap-2.0.0	dnascope-1.0-202112.01-PO	SNP	10	0.9977+-0.0003	0.9941+-0.0007	0.9959+-0.0004
			Indel	10	0.9962+-0.0005	0.9944+-0.0007	0.9953+-0.0006
hg38_T2T_masked	snap-2.0.0	octopus-0.7.4	SNP	10	0.9935+-0.0003	0.9859+-0.0010	0.9897+-0.0006
			Indel	10	0.9477+-0.0037	0.9629+-0.0045	0.9552+-0.0040
hg38_T2T_masked	snap-2.0.0	strelka-2.9.10	SNP	10	0.9839+-0.0045	0.9935+-0.0007	0.9887+-0.0025
			Indel	10	0.9855+-0.0009	0.9854+-0.0020	0.9855+-0.0014
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r11	SNP	10	0.9975+-0.0002	0.9694+-0.0012	0.9832+-0.0006
			Indel	10	0.9949+-0.0006	0.9706+-0.0018	0.9826+-0.0012
hg38_asm5_alt	dragmap-1.2.1	dnascope-1.0-202112.01-PO	SNP	10	0.9980+-0.0001	0.9730+-0.0010	0.9854+-0.0005
			Indel	10	0.9966+-0.0005	0.9776+-0.0009	0.9870+-0.0007
hg38_asm5_alt	dragmap-1.2.1	octopus-0.7.4	SNP	10	0.9968+-0.0002	0.9630+-0.0015	0.9796+-0.0007
			Indel	10	0.9497+-0.0038	0.9455+-0.0044	0.9476+-0.0040
hg38_asm5_alt	dragmap-1.2.1	strelka-2.9.10	SNP	10	0.9881+-0.0041	0.9696+-0.0010	0.9788+-0.0024
			Indel	10	0.9898+-0.0011	0.9661+-0.0021	0.9778+-0.0016
hg38_asm5_alt	sentieon-202112.01	clair3-0.1-r11	SNP	10	0.9975+-0.0003	0.9922+-0.0009	0.9949+-0.0005
			Indel	10	0.9950+-0.0005	0.9891+-0.0011	0.9920+-0.0008
hg38_asm5_alt	sentieon-202112.01	dnascope-1.0-202112.01-PO	SNP	10	0.9985+-0.0002	0.9936+-0.0007	0.9960+-0.0004
			Indel	10	0.9970+-0.0005	0.9943+-0.0008	0.9956+-0.0006
hg38_asm5_alt	sentieon-202112.01	octopus-0.7.4	SNP	10	0.9956+-0.0003	0.9845+-0.0011	0.9900+-0.0005
			Indel	10	0.9493+-0.0037	0.9627+-0.0043	0.9559+-0.0039
hg38_asm5_alt	sentieon-202112.01	strelka-2.9.10	SNP	10	0.9869+-0.0040	0.9922+-0.0007	0.9895+-0.0023
			Indel	10	0.9887+-0.0010	0.9844+-0.0020	0.9866+-0.0015
hg38_asm5_alt	sentieon-202112.01-recal	clair3-0.1-r11	SNP	10	0.9978+-0.0002	0.9911+-0.0008	0.9944+-0.0004
			Indel	10	0.9945+-0.0007	0.9882+-0.0011	0.9913+-0.0008
hg38_asm5_alt	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	SNP	10	0.9987+-0.0002	0.9929+-0.0008	0.9958+-0.0004
			Indel	10	0.9967+-0.0006	0.9925+-0.0013	0.9946+-0.0009
hg38_asm5_alt	sentieon-202112.01-recal	octopus-0.7.4	SNP	10	0.9957+-0.0003	0.9844+-0.0011	0.9900+-0.0005
			Indel	10	0.9501+-0.0036	0.9628+-0.0043	0.9564+-0.0039
hg38_asm5_alt	sentieon-202112.01-recal	strelka-2.9.10	SNP	10	0.9909+-0.0015	0.9920+-0.0007	0.9915+-0.0010

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
			Indel	10	0.9888+-0.0010	0.9844+-0.0019	0.9866+-0.0015
hg38_asm5_alt	snap-2.0.0	clair3-0.1-r11	SNP	10	0.9965+-0.0004	0.9920+-0.0007	0.9942+-0.0005
			Indel	10	0.9943+-0.0006	0.9897+-0.0014	0.9920+-0.0009
hg38_asm5_alt	snap-2.0.0	dnascope-1.0-202112.01-PO	SNP	10	0.9978+-0.0003	0.9939+-0.0007	0.9959+-0.0004
			Indel	10	0.9963+-0.0005	0.9943+-0.0007	0.9953+-0.0006
hg38_asm5_alt	snap-2.0.0	octopus-0.7.4	SNP	10	0.9941+-0.0003	0.9857+-0.0010	0.9899+-0.0005
			Indel	10	0.9481+-0.0037	0.9628+-0.0045	0.9554+-0.0040
hg38_asm5_alt	snap-2.0.0	strelka-2.9.10	SNP	10	0.9846+-0.0044	0.9933+-0.0007	0.9889+-0.0025
			Indel	10	0.9861+-0.0009	0.9853+-0.0020	0.9857+-0.0014

3.2 Preparation: "PacBio-CCS-high"

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_T2T_masked	minimap2-2.23	clair3-0.1-r11	SNP	7	0.9979+-0.0003	0.9994+-0.0001	0.9986+-0.0002
			Indel	7	0.9880+-0.0089	0.9858+-0.0063	0.9869+-0.0076
hg38_T2T_masked	minimap2-2.23	dnascope_hifi-0.4-202112.01	SNP	7	0.9993+-0.0002	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9935+-0.0049	0.9932+-0.0050	0.9934+-0.0049
hg38_T2T_masked	minimap2-2.23	peppermargindv-r0.8	SNP	7	0.9983+-0.0003	0.9990+-0.0001	0.9986+-0.0001
			Indel	7	0.9886+-0.0056	0.9858+-0.0055	0.9872+-0.0055
hg38_T2T_masked	pbmm2-1.7.0	clair3-0.1-r11	SNP	7	0.9987+-0.0002	0.9993+-0.0002	0.9990+-0.0001
			Indel	7	0.9903+-0.0087	0.9913+-0.0060	0.9908+-0.0073
hg38_T2T_masked	pbmm2-1.7.0	dnascope_hifi-0.4-202112.01	SNP	7	0.9993+-0.0003	0.9993+-0.0001	0.9994+-0.0001
			Indel	7	0.9940+-0.0046	0.9940+-0.0045	0.9940+-0.0046
hg38_T2T_masked	pbmm2-1.7.0	peppermargindv-r0.8	SNP	7	0.9989+-0.0002	0.9989+-0.0002	0.9989+-0.0001
			Indel	7	0.9908+-0.0055	0.9914+-0.0049	0.9911+-0.0052
hg38_T2T_masked	sentieon_mm2-202112.01	clair3-0.1-r11	SNP	7	0.9979+-0.0003	0.9994+-0.0001	0.9986+-0.0001
			Indel	7	0.9880+-0.0089	0.9858+-0.0063	0.9869+-0.0076
hg38_T2T_masked	sentieon_mm2-202112.01	dnascope_hifi-0.4-202112.01	SNP	7	0.9993+-0.0002	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9935+-0.0049	0.9932+-0.0050	0.9934+-0.0049

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_T2T_masked	sentieon_mm2-202112.01	peppermargindv-r0.8	SNP	7	0.9983+-0.0003	0.9990+-0.0001	0.9986+-0.0001
			Indel	7	0.9886+-0.0056	0.9858+-0.0055	0.9872+-0.0055

4 CMRG Comparison

GIAB recently released a pre-print describing Challenging Medically-Relevant Autosomal Genes (CMRG) benchmark. These regions are harder to capture than the typical GIAB benchmarks. This section contains the summarized results of the small variant analysis, specifically on the CMRG regions. For each aligner-caller pair, we ran RTG on the CMRG high confidence truth set (currently only for HG002 samples). The following results are gathered from the RTG 'summary.txt' file using the final row (i.e. "None" indicating no variants were filtered) and with the `--all-records` option enabled.

4.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
clinical	sentieon-201808.07	strelka-2.9.10	2	0.9542+-0.0031	0.9512+-0.0007	0.9527+-0.0019
hg38_T2T_masked	dragmap-1.2.1	clair3-0.1-r11	2	0.9708+-0.0001	0.9592+-0.0002	0.9650+-0.0000
hg38_T2T_masked	dragmap-1.2.1	dnascope-1.0-202112.01-PO	2	0.9751+-0.0004	0.9672+-0.0005	0.9711+-0.0005
hg38_T2T_masked	dragmap-1.2.1	octopus-0.7.4	2	0.9804+-0.0007	0.9634+-0.0009	0.9718+-0.0008
hg38_T2T_masked	dragmap-1.2.1	strelka-2.9.10	2	0.9405+-0.0033	0.9562+-0.0008	0.9482+-0.0021
hg38_T2T_masked	sentieon-202112.01	clair3-0.1-r11	2	0.9696+-0.0004	0.9594+-0.0010	0.9645+-0.0004
hg38_T2T_masked	sentieon-202112.01	dnascope-1.0-202112.01-PO	2	0.9753+-0.0011	0.9664+-0.0015	0.9708+-0.0013
hg38_T2T_masked	sentieon-202112.01	octopus-0.7.4	2	0.9734+-0.0016	0.9640+-0.0015	0.9687+-0.0016
hg38_T2T_masked	sentieon-202112.01	strelka-2.9.10	2	0.9357+-0.0045	0.9608+-0.0006	0.9481+-0.0025
hg38_T2T_masked	sentieon-202112.01-recal	clair3-0.1-r11	2	0.9707+-0.0005	0.9564+-0.0010	0.9634+-0.0003
hg38_T2T_masked	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	2	0.9760+-0.0015	0.9634+-0.0017	0.9696+-0.0015
hg38_T2T_masked	sentieon-202112.01-recal	octopus-0.7.4	2	0.9738+-0.0001	0.9635+-0.0007	0.9687+-0.0005
hg38_T2T_masked	sentieon-202112.01-recal	strelka-2.9.10	2	0.9449+-0.0026	0.9604+-0.0006	0.9526+-0.0017
hg38_T2T_masked	snap-2.0.0	clair3-0.1-r11	2	0.9644+-0.0016	0.9586+-0.0011	0.9614+-0.0014
hg38_T2T_masked	snap-2.0.0	dnascope-1.0-202112.01-PO	2	0.9707+-0.0013	0.9660+-0.0006	0.9684+-0.0010
hg38_T2T_masked	snap-2.0.0	octopus-0.7.4	2	0.9670+-0.0018	0.9652+-0.0007	0.9661+-0.0005
hg38_T2T_masked	snap-2.0.0	strelka-2.9.10	2	0.9269+-0.0055	0.9606+-0.0006	0.9435+-0.0031
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r11	2	0.9750+-0.0001	0.8495+-0.0001	0.9080+-0.0001
hg38_asm5_alt	dragmap-1.2.1	deepvariant-1.2.0	2	0.9854+-0.0003	0.8513+-0.0006	0.9134+-0.0002
hg38_asm5_alt	dragmap-1.2.1	dnascope-1.0-202112.01-PO	2	0.9778+-0.0002	0.8664+-0.0005	0.9187+-0.0004
hg38_asm5_alt	dragmap-1.2.1	octopus-0.7.4	2	0.9747+-0.0000	0.8539+-0.0012	0.9103+-0.0007
hg38_asm5_alt	dragmap-1.2.1	strelka-2.9.10	2	0.9430+-0.0029	0.8447+-0.0010	0.8911+-0.0019
hg38_asm5_alt	sentieon-202112.01	clair3-0.1-r11	2	0.9777+-0.0004	0.9517+-0.0000	0.9646+-0.0002

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_asm5_alt	sentieon-202112.01	dnascope-1.0-202112.01-PO	2	0.9831+-0.0006	0.9627+-0.0013	0.9728+-0.0010
hg38_asm5_alt	sentieon-202112.01	octopus-0.7.4	2	0.9738+-0.0010	0.9551+-0.0014	0.9644+-0.0013
hg38_asm5_alt	sentieon-202112.01	strelka-2.9.10	2	0.9449+-0.0050	0.9516+-0.0007	0.9483+-0.0028
hg38_asm5_alt	sentieon-202112.01-recal	clair3-0.1-r11	2	0.9789+-0.0002	0.9486+-0.0001	0.9635+-0.0002
hg38_asm5_alt	sentieon-202112.01-recal	dnascope-1.0-202112.01-PO	2	0.9842+-0.0013	0.9599+-0.0013	0.9719+-0.0013
hg38_asm5_alt	sentieon-202112.01-recal	octopus-0.7.4	2	0.9748+-0.0000	0.9550+-0.0013	0.9648+-0.0007
hg38_asm5_alt	sentieon-202112.01-recal	strelka-2.9.10	2	0.9542+-0.0031	0.9512+-0.0007	0.9527+-0.0019
hg38_asm5_alt	snap-2.0.0	clair3-0.1-r11	2	0.9724+-0.0010	0.9505+-0.0005	0.9612+-0.0007
hg38_asm5_alt	snap-2.0.0	dnascope-1.0-202112.01-PO	2	0.9805+-0.0007	0.9587+-0.0001	0.9695+-0.0004
hg38_asm5_alt	snap-2.0.0	octopus-0.7.4	2	0.9666+-0.0018	0.9575+-0.0011	0.9620+-0.0003
hg38_asm5_alt	snap-2.0.0	strelka-2.9.10	2	0.9335+-0.0053	0.9525+-0.0008	0.9429+-0.0031

4.2 Preparation: "PacBio-CCS-high"

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_T2T_masked	minimap2-2.23	clair3-0.1-r11	1	0.9663+-0.0000	0.9926+-0.0000	0.9792+-0.0000
hg38_T2T_masked	minimap2-2.23	dnascope_hifi-0.4-202112.01	1	0.9876+-0.0000	0.9894+-0.0000	0.9885+-0.0000
hg38_T2T_masked	minimap2-2.23	peppermargindv-r0.8	1	0.9862+-0.0000	0.9908+-0.0000	0.9885+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	clair3-0.1-r11	1	0.9683+-0.0000	0.9902+-0.0000	0.9791+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	dnascope_hifi-0.4-202112.01	1	0.9881+-0.0000	0.9889+-0.0000	0.9885+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	peppermargindv-r0.8	1	0.9806+-0.0000	0.9888+-0.0000	0.9847+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	clair3-0.1-r11	1	0.9659+-0.0000	0.9925+-0.0000	0.9790+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	dnascope_hifi-0.4-202112.01	1	0.9876+-0.0000	0.9894+-0.0000	0.9885+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	peppermargindv-r0.8	1	0.9860+-0.0000	0.9908+-0.0000	0.9884+-0.0000

5 Summarized Haplotyper Comparison

This section contains the summarized results of the haplotype analysis. For each aligner-haplotyper pair, we attempted to run the haplotyper and gather the resulting prediction. "Output" is the percentage that generated an output file. This is most commonly below 100% when a tool crashes while trying to run. "Computed" is the percentage that generated a prediction in the output file. This is most commonly below 100% when a tool successfully runs, but it does not generate a prediction in the output file (e.g. "None"). "Matching" is the percentage that generated the correct haplotype prediction based on our lookup table. These haplotyping tools often assume a particular data type, and are not designed to work under all of the conditions tested by this benchmark.

5.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Haplotyper	N	Output (%)	Computed (%)	Matching (%)
hg38_T2T_masked	dragmap-1.2.1	cyrius-1.1.1	10	100.00	80.00	80.00
hg38_T2T_masked	sentieon-202112.01	cyrius-1.1.1	10	100.00	90.00	90.00
hg38_T2T_masked	sentieon-202112.01-recal	cyrius-1.1.1	10	100.00	90.00	90.00
hg38_T2T_masked	snap-2.0.0	cyrius-1.1.1	10	100.00	90.00	90.00
hg38_asm5_alt	dragmap-1.2.1	cyrius-1.1.1	10	20.00	0.00	0.00
hg38_asm5_alt	sentieon-202112.01	cyrius-1.1.1	10	100.00	100.00	100.00
hg38_asm5_alt	sentieon-202112.01-recal	cyrius-1.1.1	10	100.00	100.00	100.00
hg38_asm5_alt	snap-2.0.0	cyrius-1.1.1	10	100.00	90.00	90.00

5.2 Preparation: "PacBio-CCS-high"

Reference	Aligner	Haplotyper	N	Output (%)	Computed (%)	Matching (%)
hg38_T2T_masked	minimap2-2.23	cyrius-1.1.1	7	100.00	28.57	14.29
hg38_T2T_masked	pbmm2-1.7.0	cyrius-1.1.1	7	100.00	42.86	28.57

Reference	Aligner	Haplotyper	N	Output (%)	Computed (%)	Matching (%)
hg38_T2T_masked	sentieon_mm2-202112.01	cyrius-1.1.1	7	100.00	28.57	14.29

6 Summarized Deletion Comparisons

This section contains the summarized results of the deletion variant analysis. For each aligner-caller pair, we ran Truvari on the corresponding truth set with these options: "`--pctsim 0.0 --sizemax 1000000000`". The following results are gathered from the Truvari "summary.txt" file.

6.1 Restricted calls

These evaluations are **RESTRICTED**, meaning that a BED file was used to limit the analysis regions, typically to "high-confidence" regions associated with the benchmark. This corresponds to the "`--includebed`" option in Truvari. As a result, benchmark sets that were not released with a corresponding BED file are excluded from this analysis. At this time, only the HG002 GIAB v0.6 benchmark is included in this analysis.

6.1.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Caller	N	precision	recall	f1
hg38_T2T_masked	dragmap-1.2.1	dysgu-1.3.10-PO	2	0.9016+-0.0233	0.7096+-0.0038	0.7940+-0.0066
hg38_T2T_masked	dragmap-1.2.1	dysgu-1.3.4-PO	2	0.9105+-0.0212	0.7073+-0.0035	0.7960+-0.0059
hg38_T2T_masked	dragmap-1.2.1	manta-1.6.0	2	0.9561+-0.0011	0.6677+-0.0188	0.7862+-0.0127
hg38_T2T_masked	sentieon-202112.01	dysgu-1.3.10-PO	2	0.9019+-0.0201	0.6656+-0.0023	0.7658+-0.0057
hg38_T2T_masked	sentieon-202112.01	dysgu-1.3.4-PO	2	0.9025+-0.0211	0.6508+-0.0014	0.7562+-0.0064
hg38_T2T_masked	sentieon-202112.01	manta-1.6.0	2	0.9529+-0.0007	0.6909+-0.0153	0.8009+-0.0101
hg38_T2T_masked	sentieon-202112.01-recal	dysgu-1.3.10-PO	2	0.8991+-0.0211	0.6657+-0.0024	0.7649+-0.0060
hg38_T2T_masked	sentieon-202112.01-recal	dysgu-1.3.4-PO	2	0.9031+-0.0172	0.6516+-0.0060	0.7568+-0.0020
hg38_T2T_masked	sentieon-202112.01-recal	manta-1.6.0	2	0.9518+-0.0009	0.6909+-0.0149	0.8005+-0.0097
hg38_T2T_masked	snap-2.0.0	dysgu-1.3.10-PO	2	0.9401+-0.0050	0.6418+-0.0104	0.7628+-0.0057
hg38_T2T_masked	snap-2.0.0	dysgu-1.3.4-PO	2	0.9420+-0.0040	0.6266+-0.0043	0.7526+-0.0018
hg38_T2T_masked	snap-2.0.0	manta-1.6.0	2	0.9490+-0.0008	0.6805+-0.0128	0.7925+-0.0084
hg38_asm5_alt	dragmap-1.2.1	dysgu-1.3.10-PO	2	0.9052+-0.0229	0.6944+-0.0032	0.7857+-0.0066
hg38_asm5_alt	dragmap-1.2.1	dysgu-1.3.4-PO	2	0.9146+-0.0191	0.6927+-0.0040	0.7882+-0.0045
hg38_asm5_alt	dragmap-1.2.1	manta-1.6.0	2	0.9616+-0.0007	0.6550+-0.0169	0.7791+-0.0117
hg38_asm5_alt	sentieon-202112.01	dysgu-1.3.10-PO	2	0.9013+-0.0195	0.6655+-0.0022	0.7655+-0.0056
hg38_asm5_alt	sentieon-202112.01	dysgu-1.3.4-PO	2	0.9061+-0.0184	0.6463+-0.0002	0.7544+-0.0062
hg38_asm5_alt	sentieon-202112.01	manta-1.6.0	2	0.9563+-0.0018	0.6899+-0.0149	0.8015+-0.0094
hg38_asm5_alt	sentieon-202112.01-recal	dysgu-1.3.10-PO	2	0.8991+-0.0206	0.6651+-0.0028	0.7645+-0.0056
hg38_asm5_alt	sentieon-202112.01-recal	dysgu-1.3.4-PO	2	0.9059+-0.0176	0.6464+-0.0042	0.7543+-0.0032
hg38_asm5_alt	sentieon-202112.01-recal	manta-1.6.0	2	0.9547+-0.0018	0.6896+-0.0140	0.8007+-0.0088

Reference	Aligner	Caller	N	precision	recall	f1
hg38_asm5_alt	snap-2.0.0	dysgu-1.3.10-PO	2	0.9411+-0.0039	0.6406+-0.0104	0.7623+-0.0061
hg38_asm5_alt	snap-2.0.0	dysgu-1.3.4-PO	2	0.9430+-0.0048	0.6225+-0.0043	0.7500+-0.0016
hg38_asm5_alt	snap-2.0.0	manta-1.6.0	2	0.9509+-0.0022	0.6791+-0.0127	0.7923+-0.0079

6.1.2 Preparation: "PacBio-CCS-high"

Reference	Aligner	Caller	N	precision	recall	f1
hg38_T2T_masked	minimap2-2.23	dysgu-1.3.10-PO	1	0.9463+-0.0000	0.9729+-0.0000	0.9594+-0.0000
hg38_T2T_masked	minimap2-2.23	dysgu-1.3.4-PO	1	0.9415+-0.0000	0.9731+-0.0000	0.9571+-0.0000
hg38_T2T_masked	minimap2-2.23	pbsv-2.8.0	1	0.9463+-0.0000	0.9847+-0.0000	0.9651+-0.0000
hg38_T2T_masked	minimap2-2.23	sniffles-2.0.2	1	0.9494+-0.0000	0.9808+-0.0000	0.9649+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	dysgu-1.3.10-PO	1	0.9372+-0.0000	0.9729+-0.0000	0.9547+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	dysgu-1.3.4-PO	1	0.9554+-0.0000	0.1849+-0.0000	0.3098+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	pbsv-2.8.0	1	0.9420+-0.0000	0.9859+-0.0000	0.9634+-0.0000
hg38_T2T_masked	pbmm2-1.7.0	sniffles-2.0.2	1	0.9450+-0.0000	0.9811+-0.0000	0.9627+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	dysgu-1.3.10-PO	1	0.9468+-0.0000	0.9736+-0.0000	0.9600+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	dysgu-1.3.4-PO	1	0.9415+-0.0000	0.9727+-0.0000	0.9568+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	pbsv-2.8.0	1	0.9459+-0.0000	0.9847+-0.0000	0.9649+-0.0000
hg38_T2T_masked	sentieon_mm2-202112.01	sniffles-2.0.2	1	0.9492+-0.0000	0.9808+-0.0000	0.9647+-0.0000

6.2 Unrestricted calls

These evaluations are **UNRESTRICTED**, meaning that there is no BED file used to limit the analysis to a subset of the genome (e.g. no high-confidence region applied). As a result, precision results are skewed downwards and should be considered a lower-bound on precision with respect to the full benchmark dataset. However, this allows us to add benchmarks that were not released with a BED file, specifically HG001/NA12878's Personalis benchmark set.

6.2.1 Preparation: "PCR-free-illumina"

Reference	Aligner	Caller	N	precision	recall	f1
hg38_T2T_masked	dragmap-1.2.1	dysgu-1.3.10-PO	4	0.5638+-0.0896	0.7866+-0.1637	0.6354+-0.0077
hg38_T2T_masked	dragmap-1.2.1	dysgu-1.3.4-PO	4	0.5682+-0.0907	0.7865+-0.1645	0.6380+-0.0072
hg38_T2T_masked	dragmap-1.2.1	manta-1.6.0	4	0.5691+-0.0968	0.7417+-0.1500	0.6221+-0.0078
hg38_T2T_masked	sentieon-202112.01	dysgu-1.3.10-PO	4	0.5717+-0.0914	0.7482+-0.1599	0.6260+-0.0068
hg38_T2T_masked	sentieon-202112.01	dysgu-1.3.4-PO	4	0.5693+-0.0916	0.7369+-0.1598	0.6199+-0.0078
hg38_T2T_masked	sentieon-202112.01	manta-1.6.0	4	0.5404+-0.0990	0.7589+-0.1444	0.6098+-0.0175
hg38_T2T_masked	sentieon-202112.01-recal	dysgu-1.3.10-PO	4	0.5714+-0.0919	0.7485+-0.1599	0.6258+-0.0064
hg38_T2T_masked	sentieon-202112.01-recal	dysgu-1.3.4-PO	4	0.5692+-0.0924	0.7371+-0.1589	0.6198+-0.0064
hg38_T2T_masked	sentieon-202112.01-recal	manta-1.6.0	4	0.5392+-0.0979	0.7588+-0.1451	0.6092+-0.0167
hg38_T2T_masked	snap-2.0.0	dysgu-1.3.10-PO	4	0.6112+-0.0957	0.7082+-0.1522	0.6336+-0.0120
hg38_T2T_masked	snap-2.0.0	dysgu-1.3.4-PO	4	0.6057+-0.0977	0.6900+-0.1455	0.6229+-0.0111
hg38_T2T_masked	snap-2.0.0	manta-1.6.0	4	0.4960+-0.0947	0.7504+-0.1473	0.5758+-0.0217
hg38_asm5_alt	dragmap-1.2.1	dysgu-1.3.10-PO	4	0.5743+-0.0933	0.7661+-0.1580	0.6348+-0.0078
hg38_asm5_alt	dragmap-1.2.1	dysgu-1.3.4-PO	4	0.5745+-0.0922	0.7648+-0.1575	0.6346+-0.0062
hg38_asm5_alt	dragmap-1.2.1	manta-1.6.0	4	0.5851+-0.1027	0.7229+-0.1443	0.6242+-0.0073
hg38_asm5_alt	sentieon-202112.01	dysgu-1.3.10-PO	4	0.5686+-0.0904	0.7475+-0.1594	0.6241+-0.0072
hg38_asm5_alt	sentieon-202112.01	dysgu-1.3.4-PO	4	0.5629+-0.0909	0.7279+-0.1555	0.6130+-0.0078
hg38_asm5_alt	sentieon-202112.01	manta-1.6.0	4	0.5502+-0.1032	0.7558+-0.1430	0.6148+-0.0182
hg38_asm5_alt	sentieon-202112.01-recal	dysgu-1.3.10-PO	4	0.5677+-0.0898	0.7477+-0.1600	0.6235+-0.0067
hg38_asm5_alt	sentieon-202112.01-recal	dysgu-1.3.4-PO	4	0.5628+-0.0919	0.7274+-0.1549	0.6126+-0.0057
hg38_asm5_alt	sentieon-202112.01-recal	manta-1.6.0	4	0.5492+-0.1020	0.7553+-0.1435	0.6141+-0.0173
hg38_asm5_alt	snap-2.0.0	dysgu-1.3.10-PO	4	0.6121+-0.0976	0.7053+-0.1509	0.6327+-0.0108
hg38_asm5_alt	snap-2.0.0	dysgu-1.3.4-PO	4	0.6027+-0.0977	0.6880+-0.1475	0.6199+-0.0122
hg38_asm5_alt	snap-2.0.0	manta-1.6.0	4	0.4999+-0.0970	0.7479+-0.1467	0.5774+-0.0223

6.2.2 Preparation: "PacBio-CCS-high"

Reference	Aligner	Caller	N	precision	recall	f1
hg38_T2T_masked	minimap2-2.23	dysgu-1.3.10-PO	2	0.4273+-0.1364	0.9429+-0.0214	0.5732+-0.1263
hg38_T2T_masked	minimap2-2.23	dysgu-1.3.4-PO	2	0.4148+-0.1336	0.9438+-0.0201	0.5618+-0.1264

Reference	Aligner	Caller	N	precision	recall	f1
hg38_T2T_masked	minimap2-2.23	pbsv-2.8.0	2	0.3681+-0.1158	0.9564+-0.0071	0.5204+-0.1206
hg38_T2T_masked	minimap2-2.23	sniffles-2.0.2	2	0.3764+-0.1182	0.9473+-0.0192	0.5265+-0.1189
hg38_T2T_masked	pbmm2-1.7.0	dysgu-1.3.10-PO	2	0.4160+-0.1336	0.9440+-0.0172	0.5632+-0.1267
hg38_T2T_masked	pbmm2-1.7.0	dysgu-1.3.4-PO	2	0.6125+-0.0607	0.4066+-0.2475	0.4280+-0.1706
hg38_T2T_masked	pbmm2-1.7.0	pbsv-2.8.0	2	0.3813+-0.1210	0.9593+-0.0087	0.5338+-0.1234
hg38_T2T_masked	pbmm2-1.7.0	sniffles-2.0.2	2	0.3836+-0.1205	0.9486+-0.0172	0.5339+-0.1203
hg38_T2T_masked	sentieon_mm2-202112.01	dysgu-1.3.10-PO	2	0.4275+-0.1363	0.9434+-0.0216	0.5736+-0.1261
hg38_T2T_masked	sentieon_mm2-202112.01	dysgu-1.3.4-PO	2	0.4148+-0.1338	0.9434+-0.0201	0.5617+-0.1265
hg38_T2T_masked	sentieon_mm2-202112.01	pbsv-2.8.0	2	0.3672+-0.1156	0.9561+-0.0070	0.5195+-0.1206
hg38_T2T_masked	sentieon_mm2-202112.01	sniffles-2.0.2	2	0.3761+-0.1180	0.9473+-0.0192	0.5262+-0.1187