

Pipeline Report

J. Matthew Holt

January 7, 2022

Contents

1	Metadata information	2
1.1	Constants	2
1.2	Dynamic	2
2	Summarized Small Variant Comparison	3
2.1	Preparation: “PCR-free-illumina”	3
2.2	Preparation: “PacBio-CCS-high”	4
3	Summarized SNP/Indel Comparison	6
3.1	Preparation: “PCR-free-illumina”	6
3.2	Preparation: “PacBio-CCS-high”	9
4	CMRG Comparison	11
4.1	Preparation: “PCR-free-illumina”	11
4.2	Preparation: “PacBio-CCS-high”	12
5	Summarized Haplotyper Comparison	14
6	Summarized Deletion Comparisons	15
6.1	Preparation: “PCR-free-illumina”	15
6.2	Preparation: “PacBio-CCS-high”	15

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

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 (≈ 3.6 M 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	8	0.9923+-0.0014	0.9920+-0.0008	0.9922+-0.0010
hg38_GIAB_masked	dragmap-1.2.1	clair3-0.1-r9	8	0.9976+-0.0003	0.9924+-0.0006	0.9950+-0.0004
hg38_GIAB_masked	dragmap-1.2.1	deepvariant-1.2.0	8	0.9989+-0.0002	0.9928+-0.0006	0.9959+-0.0004
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04	8	0.9863+-0.0007	0.9945+-0.0007	0.9904+-0.0007
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04-PO	8	0.9984+-0.0002	0.9924+-0.0007	0.9954+-0.0004
hg38_GIAB_masked	dragmap-1.2.1	sentieon_gatk4-202010.04	8	0.9913+-0.0008	0.9924+-0.0009	0.9918+-0.0008
hg38_GIAB_masked	sentieon-202010.04	clair3-0.1-r9	8	0.9974+-0.0003	0.9936+-0.0006	0.9955+-0.0004
hg38_GIAB_masked	sentieon-202010.04	deepvariant-1.2.0	8	0.9988+-0.0003	0.9937+-0.0006	0.9963+-0.0003
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04	8	0.9817+-0.0009	0.9947+-0.0006	0.9882+-0.0007
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04-PO	8	0.9980+-0.0002	0.9927+-0.0006	0.9953+-0.0003
hg38_GIAB_masked	sentieon-202010.04	octopus-0.7.4	8	0.9946+-0.0004	0.9903+-0.0008	0.9924+-0.0005
hg38_GIAB_masked	sentieon-202010.04	sentieon_gatk4-202010.04	8	0.9865+-0.0009	0.9934+-0.0007	0.9899+-0.0007
hg38_GIAB_masked	sentieon-202010.04	strelka-2.9.10	8	0.9869+-0.0039	0.9931+-0.0007	0.9900+-0.0022
hg38_GIAB_masked	sentieon-202010.04-recal	clair3-0.1-r9	8	0.9974+-0.0003	0.9926+-0.0007	0.9950+-0.0004
hg38_GIAB_masked	sentieon-202010.04-recal	deepvariant-1.2.0	8	0.9987+-0.0002	0.9935+-0.0005	0.9961+-0.0003
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04	8	0.9839+-0.0008	0.9945+-0.0007	0.9892+-0.0007
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	8	0.9983+-0.0002	0.9912+-0.0009	0.9948+-0.0005
hg38_GIAB_masked	sentieon-202010.04-recal	octopus-0.7.4	8	0.9947+-0.0004	0.9902+-0.0007	0.9925+-0.0005
hg38_GIAB_masked	sentieon-202010.04-recal	sentieon_gatk4-202010.04	8	0.9888+-0.0005	0.9931+-0.0008	0.9909+-0.0005
hg38_GIAB_masked	sentieon-202010.04-recal	strelka-2.9.10	8	0.9905+-0.0016	0.9930+-0.0007	0.9918+-0.0011
hg38_GIAB_masked	snap-2.0.0	clair3-0.1-r9	8	0.9966+-0.0003	0.9928+-0.0005	0.9947+-0.0004
hg38_GIAB_masked	snap-2.0.0	deepvariant-1.2.0	8	0.9984+-0.0002	0.9937+-0.0005	0.9960+-0.0003
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04	8	0.9820+-0.0006	0.9948+-0.0006	0.9884+-0.0006
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04-PO	8	0.9973+-0.0002	0.9932+-0.0006	0.9953+-0.0003
hg38_GIAB_masked	snap-2.0.0	sentieon_gatk4-202010.04	8	0.9849+-0.0009	0.9939+-0.0007	0.9893+-0.0007

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r9	8	0.9975+-0.0002	0.9705+-0.0008	0.9838+-0.0004
hg38_asm5_alt	dragmap-1.2.1	deepvariant-1.2.0	8	0.9986+-0.0002	0.9709+-0.0008	0.9846+-0.0004
hg38_asm5_alt	dragmap-1.2.1	dnascope-0.5-202010.04	8	0.9902+-0.0003	0.9725+-0.0009	0.9813+-0.0006
hg38_asm5_alt	dragmap-1.2.1	dnascope-0.5-202010.04-PO	8	0.9983+-0.0002	0.9710+-0.0011	0.9844+-0.0006
hg38_asm5_alt	dragmap-1.2.1	sentieon_gatk4-202010.04	8	0.9935+-0.0006	0.9702+-0.0011	0.9817+-0.0008
hg38_asm5_alt	parabricks-3.5.0	germline	1	0.9909+-0.0000	0.9944+-0.0000	0.9927+-0.0000
hg38_asm5_alt	parabricks-3.6.0	deepvariant	1	0.9986+-0.0000	0.9945+-0.0000	0.9965+-0.0000
hg38_asm5_alt	sentieon-202010.04	clair3-0.1-r9	8	0.9977+-0.0002	0.9926+-0.0007	0.9951+-0.0004
hg38_asm5_alt	sentieon-202010.04	deepvariant-1.2.0	8	0.9988+-0.0002	0.9927+-0.0006	0.9957+-0.0003
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04	8	0.9870+-0.0003	0.9937+-0.0007	0.9903+-0.0005
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04-PO	8	0.9980+-0.0002	0.9921+-0.0007	0.9950+-0.0004
hg38_asm5_alt	sentieon-202010.04	octopus-0.7.4	8	0.9958+-0.0002	0.9893+-0.0008	0.9926+-0.0004
hg38_asm5_alt	sentieon-202010.04	sentieon_gatk4-202010.04	8	0.9899+-0.0005	0.9924+-0.0007	0.9912+-0.0006
hg38_asm5_alt	sentieon-202010.04	strelka-2.9.10	8	0.9890+-0.0038	0.9921+-0.0007	0.9906+-0.0022
hg38_asm5_alt	sentieon-202010.04-recal	clair3-0.1-r9	8	0.9978+-0.0002	0.9916+-0.0007	0.9947+-0.0004
hg38_asm5_alt	sentieon-202010.04-recal	deepvariant-1.2.0	8	0.9986+-0.0002	0.9925+-0.0006	0.9956+-0.0004
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04	8	0.9892+-0.0003	0.9935+-0.0007	0.9913+-0.0004
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	8	0.9983+-0.0002	0.9907+-0.0009	0.9945+-0.0005
hg38_asm5_alt	sentieon-202010.04-recal	octopus-0.7.4	8	0.9960+-0.0002	0.9893+-0.0008	0.9926+-0.0004
hg38_asm5_alt	sentieon-202010.04-recal	sentieon_gatk4-202010.04	8	0.9922+-0.0002	0.9922+-0.0008	0.9922+-0.0004
hg38_asm5_alt	sentieon-202010.04-recal	strelka-2.9.10	8	0.9923+-0.0014	0.9920+-0.0008	0.9922+-0.0010
hg38_asm5_alt	snap-2.0.0	clair3-0.1-r9	8	0.9967+-0.0003	0.9927+-0.0005	0.9947+-0.0004
hg38_asm5_alt	snap-2.0.0	deepvariant-1.2.0	8	0.9984+-0.0002	0.9936+-0.0005	0.9960+-0.0003
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04	8	0.9843+-0.0005	0.9947+-0.0006	0.9895+-0.0005
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04-PO	8	0.9973+-0.0002	0.9932+-0.0006	0.9953+-0.0003
hg38_asm5_alt	snap-2.0.0	sentieon_gatk4-202010.04	8	0.9867+-0.0008	0.9937+-0.0007	0.9902+-0.0007

2.2 Preparation: “PacBio-CCS-high”

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	minimap2-2.23	clair3-0.1-r9	7	0.9980+-0.0013	0.9987+-0.0006	0.9983+-0.0010
hg38_GIAB_masked	minimap2-2.23	deepvariant-1.2.0	7	0.9976+-0.0018	0.9976+-0.0014	0.9976+-0.0016
hg38_GIAB_masked	minimap2-2.23	dnascope_hifi-0.4-202010.04	7	0.9986+-0.0008	0.9986+-0.0006	0.9986+-0.0007
hg38_GIAB_masked	minimap2-2.23	peppermargindv-r0.7	7	0.9983+-0.0009	0.9984+-0.0006	0.9983+-0.0007
hg38_GIAB_masked	pbmm2-1.4.0	clair3-0.1-r9	7	0.9984+-0.0013	0.9987+-0.0006	0.9985+-0.0009
hg38_GIAB_masked	pbmm2-1.4.0	deepvariant-1.2.0	7	0.9978+-0.0018	0.9975+-0.0014	0.9976+-0.0016
hg38_GIAB_masked	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	7	0.9987+-0.0008	0.9987+-0.0006	0.9987+-0.0007
hg38_GIAB_masked	pbmm2-1.4.0	peppermargindv-r0.7	7	0.9986+-0.0008	0.9983+-0.0006	0.9984+-0.0007
hg38_GIAB_masked	pbmm2-1.4.0-dvphased-1.2.0	deepvariant_hp-1.2.0	7	0.9987+-0.0009	0.9984+-0.0006	0.9985+-0.0008
hg38_GIAB_masked	sentieon_mm2-202010.04	clair3-0.1-r9	7	0.9980+-0.0013	0.9987+-0.0006	0.9983+-0.0010
hg38_GIAB_masked	sentieon_mm2-202010.04	deepvariant-1.2.0	7	0.9976+-0.0018	0.9976+-0.0014	0.9976+-0.0016
hg38_GIAB_masked	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	7	0.9986+-0.0008	0.9986+-0.0006	0.9986+-0.0007
hg38_GIAB_masked	sentieon_mm2-202010.04	peppermargindv-r0.7	7	0.9983+-0.0009	0.9984+-0.0006	0.9983+-0.0007
hg38_noalt	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	7	0.9987+-0.0008	0.9987+-0.0006	0.9987+-0.0007
hg38_noalt	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	7	0.9986+-0.0008	0.9986+-0.0006	0.9986+-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 (≈ 3.6 M 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	8	0.9912+-0.0016	0.9922+-0.0007	0.9917+-0.0010
			Indel	8	0.9885+-0.0009	0.9839+-0.0018	0.9862+-0.0013
hg38_GIAB_masked	dragmap-1.2.1	clair3-0.1-r9	SNP	8	0.9976+-0.0003	0.9925+-0.0006	0.9950+-0.0003
			Indel	8	0.9949+-0.0006	0.9896+-0.0011	0.9922+-0.0008
hg38_GIAB_masked	dragmap-1.2.1	deepvariant-1.2.0	SNP	8	0.9989+-0.0002	0.9927+-0.0007	0.9958+-0.0004
			Indel	8	0.9964+-0.0004	0.9915+-0.0008	0.9940+-0.0006
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04	SNP	8	0.9858+-0.0008	0.9944+-0.0007	0.9900+-0.0007
			Indel	8	0.9896+-0.0009	0.9945+-0.0007	0.9920+-0.0008
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04-PO	SNP	8	0.9987+-0.0002	0.9923+-0.0006	0.9955+-0.0003
			Indel	8	0.9966+-0.0004	0.9926+-0.0009	0.9946+-0.0006
hg38_GIAB_masked	dragmap-1.2.1	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_GIAB_masked	sentieon-202010.04	clair3-0.1-r9	SNP	8	0.9973+-0.0003	0.9938+-0.0007	0.9955+-0.0004
			Indel	8	0.9948+-0.0006	0.9900+-0.0012	0.9924+-0.0009
hg38_GIAB_masked	sentieon-202010.04	deepvariant-1.2.0	SNP	8	0.9988+-0.0003	0.9937+-0.0006	0.9963+-0.0003
			Indel	8	0.9963+-0.0004	0.9918+-0.0009	0.9940+-0.0006
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04	SNP	8	0.9807+-0.0009	0.9946+-0.0006	0.9876+-0.0007
			Indel	8	0.9874+-0.0010	0.9944+-0.0007	0.9909+-0.0008
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04-PO	SNP	8	0.9981+-0.0002	0.9927+-0.0006	0.9954+-0.0003
			Indel	8	0.9964+-0.0004	0.9926+-0.0008	0.9945+-0.0006
hg38_GIAB_masked	sentieon-202010.04	octopus-0.7.4	SNP	8	0.9944+-0.0004	0.9854+-0.0011	0.9899+-0.0006
			Indel	8	0.9469+-0.0026	0.9612+-0.0023	0.9540+-0.0023
hg38_GIAB_masked	sentieon-202010.04	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	sentieon-202010.04	strelka-2.9.10	SNP	8	0.9852+-0.0045	0.9934+-0.0007	0.9893+-0.0025
			Indel	8	0.9870+-0.0010	0.9844+-0.0018	0.9857+-0.0014
hg38_GIAB_masked	sentieon-202010.04-recal	clair3-0.1-r9	SNP	8	0.9974+-0.0003	0.9928+-0.0007	0.9951+-0.0004
			Indel	8	0.9942+-0.0008	0.9891+-0.0013	0.9916+-0.0010
hg38_GIAB_masked	sentieon-202010.04-recal	deepvariant-1.2.0	SNP	8	0.9986+-0.0003	0.9934+-0.0005	0.9960+-0.0003
			Indel	8	0.9963+-0.0004	0.9921+-0.0008	0.9942+-0.0006
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04	SNP	8	0.9833+-0.0009	0.9945+-0.0006	0.9888+-0.0007
			Indel	8	0.9875+-0.0010	0.9937+-0.0010	0.9906+-0.0010
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	SNP	8	0.9986+-0.0002	0.9915+-0.0007	0.9950+-0.0004
			Indel	8	0.9961+-0.0006	0.9889+-0.0021	0.9925+-0.0014
hg38_GIAB_masked	sentieon-202010.04-recal	octopus-0.7.4	SNP	8	0.9944+-0.0003	0.9854+-0.0011	0.9899+-0.0006
			Indel	8	0.9476+-0.0025	0.9613+-0.0023	0.9544+-0.0022
hg38_GIAB_masked	sentieon-202010.04-recal	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_GIAB_masked	sentieon-202010.04-recal	strelka-2.9.10	SNP	8	0.9893+-0.0018	0.9932+-0.0007	0.9913+-0.0012
			Indel	8	0.9870+-0.0010	0.9844+-0.0018	0.9857+-0.0014
hg38_GIAB_masked	snap-2.0.0	clair3-0.1-r9	SNP	8	0.9964+-0.0003	0.9928+-0.0005	0.9946+-0.0004
			Indel	8	0.9943+-0.0006	0.9903+-0.0012	0.9923+-0.0009
hg38_GIAB_masked	snap-2.0.0	deepvariant-1.2.0	SNP	8	0.9984+-0.0002	0.9936+-0.0005	0.9960+-0.0003
			Indel	8	0.9961+-0.0004	0.9921+-0.0009	0.9941+-0.0006
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04	SNP	8	0.9810+-0.0007	0.9947+-0.0006	0.9878+-0.0006
			Indel	8	0.9882+-0.0007	0.9946+-0.0007	0.9914+-0.0007
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04-PO	SNP	8	0.9974+-0.0002	0.9932+-0.0005	0.9953+-0.0003
			Indel	8	0.9961+-0.0004	0.9930+-0.0008	0.9946+-0.0006
hg38_GIAB_masked	snap-2.0.0	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r9	SNP	8	0.9975+-0.0002	0.9700+-0.0008	0.9836+-0.0004
			Indel	8	0.9949+-0.0006	0.9714+-0.0014	0.9830+-0.0010
hg38_asm5_alt	dragmap-1.2.1	deepvariant-1.2.0	SNP	8	0.9986+-0.0002	0.9703+-0.0008	0.9842+-0.0004
			Indel	8	0.9962+-0.0004	0.9732+-0.0011	0.9846+-0.0008
hg38_asm5_alt	dragmap-1.2.1	dnascope-0.5-202010.04	SNP	8	0.9897+-0.0002	0.9719+-0.0010	0.9808+-0.0006
			Indel	8	0.9932+-0.0008	0.9763+-0.0009	0.9847+-0.0008
hg38_asm5_alt	dragmap-1.2.1	dnascope-0.5-202010.04-PO	SNP	8	0.9985+-0.0002	0.9704+-0.0011	0.9842+-0.0006

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
			Indel	8	0.9967+-0.0005	0.9748+-0.0011	0.9856+-0.0008
hg38_asm5_alt	dragmap-1.2.1	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_asm5_alt	parabricks-3.5.0	germline	SNP	1	0.9905+-0.0000	0.9943+-0.0000	0.9924+-0.0000
			Indel	1	0.9929+-0.0000	0.9941+-0.0000	0.9935+-0.0000
hg38_asm5_alt	parabricks-3.6.0	deepvariant	SNP	1	0.9985+-0.0000	0.9945+-0.0000	0.9965+-0.0000
			Indel	1	0.9966+-0.0000	0.9930+-0.0000	0.9948+-0.0000
hg38_asm5_alt	sentieon-202010.04	clair3-0.1-r9	SNP	8	0.9976+-0.0002	0.9927+-0.0007	0.9952+-0.0004
			Indel	8	0.9950+-0.0005	0.9895+-0.0012	0.9922+-0.0009
hg38_asm5_alt	sentieon-202010.04	deepvariant-1.2.0	SNP	8	0.9988+-0.0002	0.9926+-0.0007	0.9957+-0.0003
			Indel	8	0.9963+-0.0004	0.9913+-0.0009	0.9938+-0.0006
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04	SNP	8	0.9862+-0.0003	0.9936+-0.0007	0.9898+-0.0005
			Indel	8	0.9918+-0.0006	0.9939+-0.0007	0.9929+-0.0006
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04-PO	SNP	8	0.9982+-0.0002	0.9920+-0.0007	0.9951+-0.0003
			Indel	8	0.9964+-0.0005	0.9923+-0.0008	0.9944+-0.0006
hg38_asm5_alt	sentieon-202010.04	octopus-0.7.4	SNP	8	0.9957+-0.0002	0.9844+-0.0012	0.9900+-0.0006
			Indel	8	0.9478+-0.0025	0.9608+-0.0023	0.9543+-0.0023
hg38_asm5_alt	sentieon-202010.04	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_asm5_alt	sentieon-202010.04	strelka-2.9.10	SNP	8	0.9874+-0.0043	0.9924+-0.0007	0.9899+-0.0024
			Indel	8	0.9884+-0.0008	0.9839+-0.0018	0.9861+-0.0013
hg38_asm5_alt	sentieon-202010.04-recal	clair3-0.1-r9	SNP	8	0.9978+-0.0002	0.9918+-0.0007	0.9948+-0.0004
			Indel	8	0.9945+-0.0008	0.9887+-0.0013	0.9916+-0.0010
hg38_asm5_alt	sentieon-202010.04-recal	deepvariant-1.2.0	SNP	8	0.9986+-0.0002	0.9924+-0.0006	0.9955+-0.0004
			Indel	8	0.9963+-0.0004	0.9917+-0.0008	0.9940+-0.0006
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04	SNP	8	0.9887+-0.0003	0.9935+-0.0007	0.9911+-0.0004
			Indel	8	0.9917+-0.0007	0.9931+-0.0010	0.9925+-0.0008
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	SNP	8	0.9986+-0.0001	0.9909+-0.0008	0.9947+-0.0004
			Indel	8	0.9961+-0.0007	0.9886+-0.0020	0.9923+-0.0014
hg38_asm5_alt	sentieon-202010.04-recal	octopus-0.7.4	SNP	8	0.9958+-0.0002	0.9844+-0.0012	0.9901+-0.0006
			Indel	8	0.9486+-0.0024	0.9609+-0.0023	0.9547+-0.0022
hg38_asm5_alt	sentieon-202010.04-recal	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan
hg38_asm5_alt	sentieon-202010.04-recal	strelka-2.9.10	SNP	8	0.9912+-0.0016	0.9922+-0.0007	0.9917+-0.0010

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
			Indel	8	0.9885+-0.0009	0.9839+-0.0018	0.9862+-0.0013
hg38_asm5_alt	snap-2.0.0	clair3-0.1-r9	SNP	8	0.9966+-0.0003	0.9927+-0.0005	0.9946+-0.0003
			Indel	8	0.9944+-0.0006	0.9902+-0.0011	0.9923+-0.0008
hg38_asm5_alt	snap-2.0.0	deepvariant-1.2.0	SNP	8	0.9984+-0.0002	0.9936+-0.0005	0.9960+-0.0003
			Indel	8	0.9961+-0.0004	0.9920+-0.0009	0.9940+-0.0006
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04	SNP	8	0.9834+-0.0005	0.9946+-0.0006	0.9889+-0.0005
			Indel	8	0.9899+-0.0007	0.9945+-0.0006	0.9922+-0.0007
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04-PO	SNP	8	0.9974+-0.0002	0.9932+-0.0006	0.9953+-0.0003
			Indel	8	0.9960+-0.0004	0.9930+-0.0008	0.9946+-0.0006
hg38_asm5_alt	snap-2.0.0	sentieon_gatk4-202010.04	SNP	8	nan+-nan	0.0000+-0.0000	nan+-nan
			Indel	8	nan+-nan	0.0000+-0.0000	nan+-nan

3.2 Preparation: “PacBio-CCS-high”

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	minimap2-2.23	clair3-0.1-r9	SNP	7	0.9978+-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_GIAB_masked	minimap2-2.23	deepvariant-1.2.0	SNP	7	0.9984+-0.0003	0.9991+-0.0001	0.9988+-0.0001
			Indel	7	0.9822+-0.0127	0.9787+-0.0116	0.9805+-0.0121
hg38_GIAB_masked	minimap2-2.23	dnascope_hifi-0.4-202010.04	SNP	7	0.9993+-0.0003	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9935+-0.0049	0.9932+-0.0050	0.9934+-0.0049
hg38_GIAB_masked	minimap2-2.23	peppermargindv-r0.7	SNP	7	0.9981+-0.0002	0.9991+-0.0002	0.9986+-0.0001
			Indel	7	0.9896+-0.0057	0.9855+-0.0059	0.9875+-0.0058
hg38_GIAB_masked	pbmm2-1.4.0	clair3-0.1-r9	SNP	7	0.9987+-0.0002	0.9993+-0.0002	0.9990+-0.0001
			Indel	7	0.9903+-0.0087	0.9912+-0.0059	0.9908+-0.0073
hg38_GIAB_masked	pbmm2-1.4.0	deepvariant-1.2.0	SNP	7	0.9990+-0.0002	0.9990+-0.0002	0.9990+-0.0001
			Indel	7	0.9846+-0.0126	0.9843+-0.0113	0.9844+-0.0119
hg38_GIAB_masked	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	SNP	7	0.9994+-0.0003	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9940+-0.0046	0.9940+-0.0045	0.9940+-0.0046

Reference	Aligner	Caller	Type	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	pbmm2-1.4.0	peppermargindv-r0.7	SNP	7	0.9989+-0.0002	0.9990+-0.0002	0.9989+-0.0001
			Indel	7	0.9918+-0.0056	0.9908+-0.0056	0.9913+-0.0056
hg38_GIAB_masked	pbmm2-1.4.0-dvphased-1.2.0	deepvariant_hp-1.2.0	SNP	7	0.9990+-0.0001	0.9990+-0.0002	0.9990+-0.0001
			Indel	7	0.9914+-0.0059	0.9910+-0.0057	0.9912+-0.0058
hg38_GIAB_masked	sentieon_mm2-202010.04	clair3-0.1-r9	SNP	7	0.9979+-0.0003	0.9994+-0.0002	0.9986+-0.0002
			Indel	7	0.9880+-0.0088	0.9859+-0.0063	0.9870+-0.0075
hg38_GIAB_masked	sentieon_mm2-202010.04	deepvariant-1.2.0	SNP	7	0.9984+-0.0003	0.9991+-0.0001	0.9988+-0.0001
			Indel	7	0.9822+-0.0127	0.9787+-0.0116	0.9805+-0.0121
hg38_GIAB_masked	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	SNP	7	0.9993+-0.0003	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9935+-0.0049	0.9932+-0.0050	0.9934+-0.0049
hg38_GIAB_masked	sentieon_mm2-202010.04	peppermargindv-r0.7	SNP	7	0.9981+-0.0002	0.9991+-0.0002	0.9986+-0.0001
			Indel	7	0.9895+-0.0057	0.9855+-0.0060	0.9875+-0.0058
hg38_noalt	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	SNP	7	0.9994+-0.0003	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9940+-0.0046	0.9940+-0.0045	0.9940+-0.0046
hg38_noalt	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	SNP	7	0.9993+-0.0003	0.9993+-0.0001	0.9993+-0.0001
			Indel	7	0.9935+-0.0049	0.9932+-0.0049	0.9934+-0.0049

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_GIAB_masked	dragmap-1.2.1	clair3-0.1-r9	2	0.9708+-0.0001	0.9596+-0.0002	0.9652+-0.0002
hg38_GIAB_masked	dragmap-1.2.1	deepvariant-1.2.0	2	0.9878+-0.0004	0.9601+-0.0005	0.9738+-0.0000
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04	2	0.9353+-0.0011	0.9659+-0.0008	0.9504+-0.0003
hg38_GIAB_masked	dragmap-1.2.1	dnascope-0.5-202010.04-PO	2	0.9876+-0.0005	0.9616+-0.0010	0.9744+-0.0007
hg38_GIAB_masked	dragmap-1.2.1	sentieon_gatk4-202010.04	2	0.9514+-0.0025	0.9527+-0.0008	0.9521+-0.0016
hg38_GIAB_masked	sentieon-202010.04	clair3-0.1-r9	2	0.9696+-0.0005	0.9604+-0.0001	0.9650+-0.0002
hg38_GIAB_masked	sentieon-202010.04	deepvariant-1.2.0	2	0.9870+-0.0002	0.9617+-0.0001	0.9743+-0.0001
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04	2	0.9244+-0.0016	0.9657+-0.0009	0.9446+-0.0004
hg38_GIAB_masked	sentieon-202010.04	dnascope-0.5-202010.04-PO	2	0.9848+-0.0002	0.9615+-0.0008	0.9730+-0.0005
hg38_GIAB_masked	sentieon-202010.04	octopus-0.7.4	2	0.9735+-0.0017	0.9640+-0.0015	0.9687+-0.0016
hg38_GIAB_masked	sentieon-202010.04	sentieon_gatk4-202010.04	2	0.9406+-0.0014	0.9567+-0.0010	0.9486+-0.0012
hg38_GIAB_masked	sentieon-202010.04	strelka-2.9.10	2	0.9357+-0.0045	0.9607+-0.0005	0.9481+-0.0025
hg38_GIAB_masked	sentieon-202010.04-recal	clair3-0.1-r9	2	0.9707+-0.0005	0.9574+-0.0000	0.9640+-0.0002
hg38_GIAB_masked	sentieon-202010.04-recal	deepvariant-1.2.0	2	0.9885+-0.0007	0.9614+-0.0005	0.9748+-0.0005
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04	2	0.9309+-0.0019	0.9648+-0.0007	0.9476+-0.0007
hg38_GIAB_masked	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	2	0.9876+-0.0005	0.9569+-0.0008	0.9720+-0.0006
hg38_GIAB_masked	sentieon-202010.04-recal	octopus-0.7.4	2	0.9738+-0.0000	0.9634+-0.0007	0.9686+-0.0004
hg38_GIAB_masked	sentieon-202010.04-recal	sentieon_gatk4-202010.04	2	0.9477+-0.0002	0.9560+-0.0007	0.9518+-0.0002
hg38_GIAB_masked	sentieon-202010.04-recal	strelka-2.9.10	2	0.9451+-0.0026	0.9604+-0.0006	0.9526+-0.0017
hg38_GIAB_masked	snap-2.0.0	clair3-0.1-r9	2	0.9645+-0.0015	0.9601+-0.0002	0.9623+-0.0008
hg38_GIAB_masked	snap-2.0.0	deepvariant-1.2.0	2	0.9869+-0.0009	0.9615+-0.0003	0.9741+-0.0003
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04	2	0.9212+-0.0009	0.9660+-0.0008	0.9430+-0.0001

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	snap-2.0.0	dnascope-0.5-202010.04-PO	2	0.9830+-0.0010	0.9629+-0.0008	0.9729+-0.0009
hg38_GIAB_masked	snap-2.0.0	sentieon_gatk4-202010.04	2	0.9343+-0.0011	0.9575+-0.0001	0.9457+-0.0007
hg38_asm5_alt	dragmap-1.2.1	clair3-0.1-r9	2	0.9750+-0.0001	0.8511+-0.0002	0.9088+-0.0002
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-0.5-202010.04	2	0.9408+-0.0021	0.8565+-0.0008	0.8967+-0.0005
hg38_asm5_alt	dragmap-1.2.1	dnascope-0.5-202010.04-PO	2	0.9827+-0.0004	0.8520+-0.0012	0.9126+-0.0008
hg38_asm5_alt	dragmap-1.2.1	sentieon_gatk4-202010.04	2	0.9587+-0.0018	0.8417+-0.0010	0.8964+-0.0014
hg38_asm5_alt	sentieon-202010.04	clair3-0.1-r9	2	0.9777+-0.0002	0.9517+-0.0000	0.9646+-0.0001
hg38_asm5_alt	sentieon-202010.04	deepvariant-1.2.0	2	0.9882+-0.0006	0.9528+-0.0002	0.9702+-0.0002
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04	2	0.9382+-0.0024	0.9567+-0.0005	0.9474+-0.0010
hg38_asm5_alt	sentieon-202010.04	dnascope-0.5-202010.04-PO	2	0.9835+-0.0004	0.9527+-0.0005	0.9679+-0.0004
hg38_asm5_alt	sentieon-202010.04	octopus-0.7.4	2	0.9738+-0.0010	0.9551+-0.0014	0.9643+-0.0012
hg38_asm5_alt	sentieon-202010.04	sentieon_gatk4-202010.04	2	0.9512+-0.0013	0.9480+-0.0014	0.9496+-0.0013
hg38_asm5_alt	sentieon-202010.04	strelka-2.9.10	2	0.9449+-0.0050	0.9516+-0.0007	0.9483+-0.0028
hg38_asm5_alt	sentieon-202010.04-recal	clair3-0.1-r9	2	0.9789+-0.0002	0.9486+-0.0000	0.9636+-0.0001
hg38_asm5_alt	sentieon-202010.04-recal	deepvariant-1.2.0	2	0.9884+-0.0009	0.9526+-0.0003	0.9702+-0.0006
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04	2	0.9447+-0.0024	0.9559+-0.0006	0.9503+-0.0008
hg38_asm5_alt	sentieon-202010.04-recal	dnascope-0.5-202010.04-PO	2	0.9859+-0.0007	0.9484+-0.0009	0.9667+-0.0008
hg38_asm5_alt	sentieon-202010.04-recal	octopus-0.7.4	2	0.9748+-0.0001	0.9550+-0.0013	0.9648+-0.0007
hg38_asm5_alt	sentieon-202010.04-recal	sentieon_gatk4-202010.04	2	0.9587+-0.0002	0.9471+-0.0008	0.9528+-0.0005
hg38_asm5_alt	sentieon-202010.04-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-r9	2	0.9724+-0.0010	0.9520+-0.0002	0.9621+-0.0006
hg38_asm5_alt	snap-2.0.0	deepvariant-1.2.0	2	0.9869+-0.0007	0.9533+-0.0004	0.9698+-0.0001
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04	2	0.9318+-0.0014	0.9578+-0.0008	0.9446+-0.0003
hg38_asm5_alt	snap-2.0.0	dnascope-0.5-202010.04-PO	2	0.9825+-0.0009	0.9546+-0.0008	0.9684+-0.0008
hg38_asm5_alt	snap-2.0.0	sentieon_gatk4-202010.04	2	0.9431+-0.0007	0.9493+-0.0003	0.9462+-0.0005

4.2 Preparation: “PacBio-CCS-high”

Reference	Aligner	Caller	N	Precision	Sensitivity	F-measure
hg38_GIAB_masked	minimap2-2.23	clair3-0.1-r9	1	0.9668+-0.0000	0.9925+-0.0000	0.9795+-0.0000
hg38_GIAB_masked	minimap2-2.23	deepvariant-1.2.0	1	0.9792+-0.0000	0.9886+-0.0000	0.9839+-0.0000
hg38_GIAB_masked	minimap2-2.23	dnascope_hifi-0.4-202010.04	1	0.9876+-0.0000	0.9894+-0.0000	0.9885+-0.0000
hg38_GIAB_masked	minimap2-2.23	peppermargindv-r0.7	1	0.9854+-0.0000	0.9908+-0.0000	0.9881+-0.0000
hg38_GIAB_masked	pbmm2-1.4.0	clair3-0.1-r9	1	0.9680+-0.0000	0.9904+-0.0000	0.9790+-0.0000
hg38_GIAB_masked	pbmm2-1.4.0	deepvariant-1.2.0	1	0.9806+-0.0000	0.9874+-0.0000	0.9840+-0.0000
hg38_GIAB_masked	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	1	0.9882+-0.0000	0.9889+-0.0000	0.9886+-0.0000
hg38_GIAB_masked	pbmm2-1.4.0	peppermargindv-r0.7	1	0.9824+-0.0000	0.9889+-0.0000	0.9856+-0.0000
hg38_GIAB_masked	pbmm2-1.4.0-dvphased-1.2.0	deepvariant_hp-1.2.0	1	0.9718+-0.0000	0.9892+-0.0000	0.9805+-0.0000
hg38_GIAB_masked	sentieon_mm2-202010.04	clair3-0.1-r9	1	0.9671+-0.0000	0.9927+-0.0000	0.9797+-0.0000
hg38_GIAB_masked	sentieon_mm2-202010.04	deepvariant-1.2.0	1	0.9792+-0.0000	0.9886+-0.0000	0.9839+-0.0000
hg38_GIAB_masked	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	1	0.9876+-0.0000	0.9894+-0.0000	0.9885+-0.0000
hg38_GIAB_masked	sentieon_mm2-202010.04	peppermargindv-r0.7	1	0.9854+-0.0000	0.9908+-0.0000	0.9881+-0.0000
hg38_noalt	pbmm2-1.4.0	dnascope_hifi-0.4-202010.04	1	0.9864+-0.0000	0.9841+-0.0000	0.9853+-0.0000
hg38_noalt	sentieon_mm2-202010.04	dnascope_hifi-0.4-202010.04	1	0.9857+-0.0000	0.9845+-0.0000	0.9851+-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. We then calculated the fraction of haplotypers that returned *a* result and the fraction that returned the *correct* result based on our lookup dictionary.

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 Personalis (related to Genome in a Bottle) high confidence deletion truth set ($\approx 2.6K$ deletions). The following results are gathered from the Truvari “summary.txt” file.

6.1 Preparation: “PCR-free-illumina”

Reference	Aligner	Caller	N	precision	recall	f1
hg38_GIAB_masked	dragmap-1.2.1	manta-1.6.0	4	0.5694+-0.0968	0.7414+-0.1496	0.6223+-0.0078
hg38_GIAB_masked	sentieon-202010.04	manta-1.6.0	4	0.5405+-0.0989	0.7585+-0.1441	0.6098+-0.0173
hg38_GIAB_masked	sentieon-202010.04-recal	manta-1.6.0	4	0.5395+-0.0979	0.7583+-0.1448	0.6092+-0.0166
hg38_GIAB_masked	snap-2.0.0	manta-1.6.0	4	0.4962+-0.0946	0.7500+-0.1470	0.5759+-0.0216
hg38_asm5_alt	dragmap-1.2.1	manta-1.6.0	4	0.5851+-0.1027	0.7227+-0.1442	0.6242+-0.0073
hg38_asm5_alt	sentieon-202010.04	manta-1.6.0	4	0.5502+-0.1032	0.7555+-0.1427	0.6147+-0.0183
hg38_asm5_alt	sentieon-202010.04-recal	manta-1.6.0	4	0.5492+-0.1020	0.7550+-0.1432	0.6140+-0.0174
hg38_asm5_alt	snap-2.0.0	manta-1.6.0	4	0.4999+-0.0970	0.7476+-0.1463	0.5773+-0.0224

6.2 Preparation: “PacBio-CCS-high”

Reference	Aligner	Caller	N	precision	recall	f1
hg38_GIAB_masked	minimap2-2.23	pbsv-2.6.2	2	0.3513+-0.1132	0.9563+-0.0069	0.5029+-0.1210
hg38_GIAB_masked	pbbmm2-1.4.0	pbsv-2.6.2	2	0.3720+-0.1221	0.9597+-0.0088	0.5238+-0.1265
hg38_GIAB_masked	sentieon_mm2-202010.04	pbsv-2.6.2	2	0.3513+-0.1137	0.9562+-0.0071	0.5027+-0.1215
hg38_noalt	pbbmm2-1.4.0	pbsv-2.6.2	2	0.3718+-0.1222	0.9597+-0.0091	0.5235+-0.1266
hg38_noalt	sentieon_mm2-202010.04	pbsv-2.6.2	2	0.3515+-0.1139	0.9563+-0.0069	0.5029+-0.1217